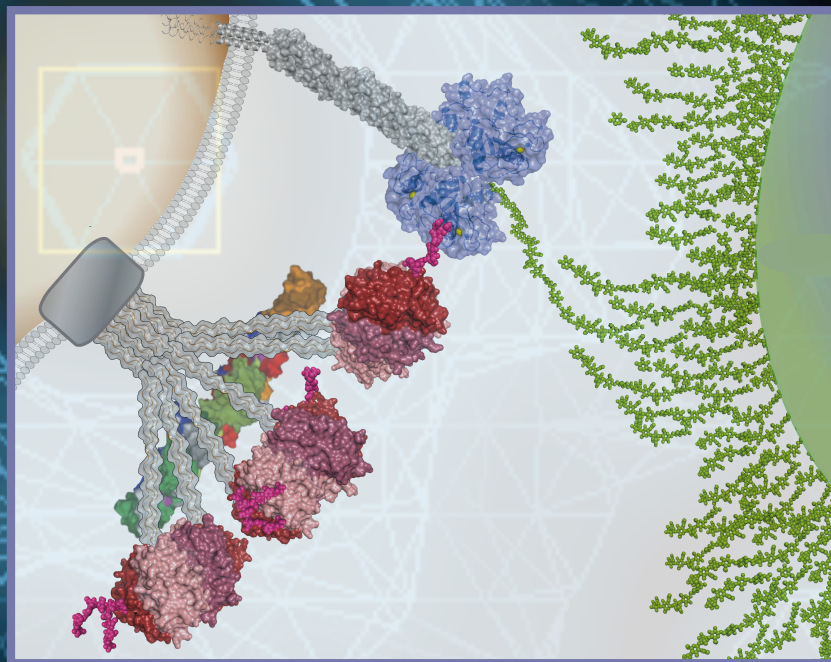


UNIVERSIDAD AUTÓNOMA DE MADRID
DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Biología estructural de las interacciones
patógeno – hospedador en *Streptococcus pneumoniae*:
mecanismos de reconocimiento y
lisis de la pared bacteriana



TESIS DOCTORAL: Noella Silva Martin

2011

DEPARTAMENTO BIOLOGÍA MOLECULAR
FACULTAD DE CIENCIA
UNIVERSIDAD AUTÓNOMA DE MADRID

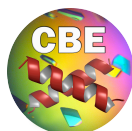


Biología Estructural de las Interacciones
Patógeno – Hospedador en *Streptococcus pneumoniae*:
Mecanismos de Reconocimiento y Lisis
de la Pared Bacteriana

Memoria que presenta la licenciada en Biología NOELLA A. R.
SILVA MARTIN para optar al grado de Doctor por la
Universidad Autónoma de Madrid

DIRECTOR: JUAN A. HERMOSO DOMÍNGUEZ

MADRID 2011



CSIC

Departamento de Cristalografía y Biología Estructural
Instituto de Química-Física Rocasolano. CSIC

Esta tesis doctoral se ha realizado gracias al disfrute de una beca predoctoral concedida a la estudiante Noella A. R. Silva Martin financiada por la Fundación Gran Mariscal de Ayacucho – Venezuela.

Esta tesis doctoral se ha realizado bajo la financiación asociada al proyecto “Biología estructural de las interacciones hospedador – patógeno en *Streptococcus pneumoniae*” (referencia BFU2008-01711/BMC), concedido al Dr. Juan A. Hermoso Domínguez y financiada por el Ministerio de Ciencia e Innovación.

RESUMEN

Streptococcus pneumoniae es el causante de enfermedades de gran impacto socioeconómico, como la neumonía, la mayor responsable de la mortalidad en niños menores de 5 años. Por ello, es de vital importancia el estudio de las relaciones patógeno hospedador que puedan ayudar a entender la biología de este patógeno y las estrategias creadas por el hospedador para defenderse de éste con el propósito de encontrar nuevas estrategias terapéuticas.

En este trabajo se ha abordado el estudio estructural y funcional de distintos sistemas proteicos involucrados en el reconocimiento de motivos moleculares de la superficie bacteriana de *S. pneumoniae*, utilizando técnicas cristalográficas. En particular esta trabajo se ha centrado en la determinación estructural de la endolisina Cpl-7, del receptor SIGN-R1 y de un anticuerpo sialilado. Todas estas proteínas están implicadas en procesos de lisis bacteriana bien sea de forma directa o indirecta.

Cpl-7 es una endolisina codificada por el bacteriófago Cp-7, la cual tiene importantes implicaciones como enzibiótico, debido a su efectividad para degradar la pared bacteriana. Se ha producido, cristalizado y determinado las estructuras cristalográficas del módulo catalítico y del dominio que forma parte del módulo de anclaje a la pared bacteriana. Este estudio ha proporcionado las bases estructurales sobre el arreglo modular de la proteína completa así como también los mecanismos catalíticos de esta endolisina.

SIGN-R1 es un receptor de membrana ubicado en la superficie de los macrófagos implicado en varios procesos como el reconocimiento de los patógenos, la activación de la respuesta inmune e inhibición de la respuesta inflamatoria. En este trabajo se han determinado las estructuras del dominio de unión a carbohidratos de SIGN-R1 así como sus complejos con sulfodextrano y con ácido siálico. Se han identificado dos sitios de unión en SIGN-R1; el primario canónico que está mediado por calcio y el secundario identificado por primera vez en esta familia de receptores independiente de calcio. En este trabajo se ha demostrado que el ácido siálico es reconocido por el sitio primario lo que, unido al trabajo estructural con el anticuerpo explica por qué SIGN-R1 reconoce sólo las proteínas del sistema inmune (Fc, C1q) sialiladas (2,6) y aclara el papel de SIGN-R1 en la modulación de la respuesta inflamatoria. La identificación de dos sitios de unión a carbohidratos en SIGN-R1 ha permitido también explicar el mecanismo de activación de la vía del complemento mediada por este receptor. En efecto, mientras el patógeno sería identificado a través del sitio secundario, la proteína del complemento (C1q) sería reconocida a través del sitio primario. De este modo SIGN-R1 sería capaz de unir simultáneamente al patógeno y al sistema del complemento para así eliminar de una manera muy eficiente a patógenos encapsulados como el neumococo.

ABSTRACT

Streptococcus pneumoniae is a major cause of life-threatening infections such as pneumonia, meningitis and septicemia. The disease burden is high both in developed and developing countries, and the high-risk groups include children, elderly persons and immunocompromised patients. Approximately one million children under 5 years of age die from pneumococcal infections every year. A prerequisite for developing new control strategies is the detailed understanding of the interaction between pathogen and the host cells.

This work addresses the study of the structural biology of some proteins that are critical in pathogen binding and lysis. Experimental approaches include the use of protein engineering, X-ray diffraction techniques and biophysical characterization. Emphasis is placed on studies into molecular aspects of the interactions between pathogen and host macromolecules. In particular, this work is focused on the structural determination of the endolysin Cpl-7, SIGN-R1 receptor and sialylated antibody.

Cpl-7 is an endolysin (phage-encoded enzyme that break down bacterial peptidoglycan) with potential application as antibacterial agents (enzybiotics). The crystal structures of catalytic module and that of the repeated domain building the cell-wall binding module have been solved. This study provides insights on the modular arrangement of the complete protein as well as on the catalytic mechanism of this endolysin.

SIGN-R1 is an important membrane receptor of macrophages responsible for the pathogen recognition and activation of the immune response as well as in antiinflammatory mechanisms. We report herein the high-resolution crystal structure of the carbohydrate recognition domain (CRD) of SIGN-R1 as well as their complexes with dextran sulphate and sialic acid. In addition, docking experiments with models of serotype 14 pneumococcal CPS and teichoic acids from *Staphylococcus aureus* have been performed. As observed in the crystal structure, the nature of sialic acid recognition by the classical lectin site of SIGN-R1 precludes the interaction with non α -2,6-sialylated glycoproteins thereby implicating the role of SIGN-R1 in anti-inflammatory activity of intravenous Ig. We found that sialylated-glycoproteins, such as C1q and Ig, bind through the classical lectin polysaccharide recognition site, while microbial pathogens are also recognised through a here described novel substrate binding site in SIGN-R1. These two different recognition sites allow SIGN-R1 to simultaneously bind both immune glycoproteins and microbial components, thereby accommodating SIGN-R1's unusual C3 activation pathway.

| | | |
|----------|--|-----------|
| 1 | INTRODUCCIÓN | 1 |
| 1.1 | <i>STREPTOCOCCUS PNEUMONIAE</i> | 1 |
| 1.1.1 | <i>Características generales</i> | 1 |
| 1.1.2 | <i>La cápsula polisacárida</i> | 2 |
| 1.1.3 | <i>La pared celular</i> | 3 |
| 1.1.3.1 | <i>El peptidoglicano</i> | 4 |
| 1.1.3.2 | <i>Ácido teicoico y otros componentes de la pared celular</i> | 5 |
| 1.1.4 | <i>Proteínas de superficie en S. pneumoniae</i> | 7 |
| 1.1.5 | <i>Enzimas líticas de pared de neumococo y sus fagos: Cpl-7 un caso especial</i> | 7 |
| 1.1.5.1 | <i>Mecanismo de hidrólisis de las lisozimas de la familia GH-25</i> | 9 |
| 1.1.6 | <i>Terapia fágica y enzibióticos</i> | 14 |
| 1.2 | ESTRATEGIAS DE RECONOCIMIENTO DEL PATÓGENO POR EL HOSPEDADOR | 16 |
| 1.2.1 | <i>Sistema inmune</i> | 16 |
| 1.2.2 | <i>Sistema del complemento</i> | 17 |
| 1.2.3 | <i>La vía clásica del complemento</i> | 19 |
| 1.2.4 | <i>Vía clásica del complemento iniciada por SIGN-R1</i> | 20 |
| 1.2.5 | <i>Familia de las lectinas tipo-C</i> | 21 |
| 1.2.6 | <i>SIGN-R1 y su participación en la inhibición del proceso inflamatorio</i> | 22 |
| 2 | OBJETIVOS | 27 |
| 3 | TÉCNICAS EXPERIMENTALES..... | 31 |
| 3.1 | <i>VISIÓN GENERAL DEL PROCESO EN CRISTALOGRAFÍA DE RAYOS X</i> | 31 |
| 3.2 | CONSIDERACIONES TEÓRICAS | 32 |
| 3.2.1 | <i>Cristalización</i> | 32 |
| 3.2.2 | <i>Principios básicos de la cristalización</i> | 33 |
| 3.2.3 | <i>Métodos de cristalización</i> | 34 |
| 3.2.4 | <i>Búsqueda de las condiciones de cristalización</i> | 35 |
| 3.2.5 | <i>Cristalización de complejos proteína-ligando</i> | 36 |
| 3.3 | TÉCNICAS CRISTALOGRÁFICAS Y RESOLUCIÓN ESTRUCTURAL | 36 |
| 3.3.1 | <i>Medidas del espectro de difracción</i> | 37 |
| 3.3.2 | <i>Fuentes de radiación y equipos de difracción</i> | 38 |
| 3.3.3 | <i>Preparación de las muestras</i> | 40 |
| 3.3.4 | <i>Procesado de los datos de difracción</i> | 40 |
| 3.3.5 | <i>Resolución estructural</i> | 41 |
| 3.4 | ENSAYOS DE “DOCKING” CON HERRAMIENTAS INFORMÁTICAS | 44 |
| 3.4.1 | <i>Programa GOLD</i> | 44 |
| 3.5 | OTROS PROGRAMAS UTILIZADOS | 45 |
| 3.6 | TÉCNICAS COMPLEMENTARIAS UTILIZADAS | 45 |
| 3.6.1 | <i>Espectrometría de masas</i> | 45 |

| | | |
|---------|---|----|
| 3.6.2 | <i>Secuenciación de proteínas</i> | 45 |
| 3.7 | RESOLUCIÓN ESTRUCTURAL DE LA ENDOLISINA CPL-7 | 46 |
| 3.7.1 | <i>Obtención de las muestras</i> | 46 |
| 3.7.2 | <i>Cristalización de Cpl-7</i> | 48 |
| 3.7.3 | <i>Resolución estructural de Cpl-7</i> | 48 |
| 3.8 | RESOLUCIÓN ESTRUCTURAL DE SIGN-R1 | 49 |
| 3.8.1 | <i>Obtención de las muestras</i> | 49 |
| 3.8.2 | <i>Cristalización del ECD_ SIGN-R1 y CRD_ SIGN-R1</i> | 50 |
| 3.8.3 | <i>Cristalización de complejos CRD_ SIGN-R1 : ligando</i> | 50 |
| 3.8.4 | <i>Determinación estructural de CRD_ SIGN-R1 y sus complejos</i> | 51 |
| 3.8.5 | <i>Ensayos de "docking" CRD_ SIGN-R1 con ribitol fosfato y cápsula polisacárida</i> | 52 |
| 3.9 | DETERMINACIÓN ESTRUCTURAL DE LA PORCIÓN Fc DE IgG | 53 |
| 3.9.1 | <i>Obtención de las muestras</i> | 53 |
| 3.9.2 | <i>Cristalización</i> | 53 |
| 3.9.3 | <i>Resolución estructural</i> | 53 |
| 3.9.4 | <i>Ensayos de dinámica molecular del azúcar terminal</i> | 54 |
| 4 | RESULTADOS | 57 |
| 4.1 | DETERMINACIÓN ESTRUCTURAL DE LA ENDOLISINA CPL-7 | 57 |
| 4.1.1 | <i>Resolución estructural del módulo catalítico de Cpl-7</i> | 57 |
| 4.1.1.1 | <i>Expresión y purificación de Cpl-7</i> | 57 |
| 4.1.1.2 | <i>Cristalización de Cpl-7</i> | 58 |
| 4.1.1.3 | <i>Integridad de la proteína Cpl-7 en los cristales</i> | 58 |
| 4.1.1.4 | <i>Medida de los datos de difracción de Cpl-7</i> | 59 |
| 4.1.1.5 | <i>Resolución estructural y refinamiento del módulo catalítico de Cpl-7</i> | 60 |
| 4.1.1.6 | <i>Análisis de la estructura tridimensional del módulo catalítico de Cpl-7</i> | 60 |
| 4.1.2 | <i>Determinación estructural del módulo de unión a la pared de Cpl-7</i> | 63 |
| 4.1.2.1 | <i>Expresión y purificación del extremo C-terminal de Cpl-7</i> | 63 |
| 4.1.2.2 | <i>Cristalización del módulo de unión a la pared de Cpl-7</i> | 64 |
| 4.1.2.3 | <i>Medida de los datos de difracción</i> | 65 |
| 4.1.2.4 | <i>Resolución estructural del módulo de unión a pared de Cpl-7</i> | 66 |
| 4.1.2.5 | <i>Análisis de la estructura tridimensional del módulo de unión a la pared de Cpl-7</i> | 67 |
| 4.2 | DETERMINACIÓN ESTRUCTURAL DEL RECEPTOR SIGN-R1 | 71 |
| 4.2.1 | <i>Cristalización de CRD_ SIGN-R1</i> | 71 |
| 4.2.2 | <i>Medida de los datos de difracción</i> | 72 |
| 4.2.3 | <i>Resolución estructural y refinamiento</i> | 72 |
| 4.2.4 | <i>Complejos cristalográficos con ligandos</i> | 73 |
| 4.2.4.1 | <i>Cristalización de los complejos</i> | 73 |
| 4.2.4.2 | <i>Medida de los datos de difracción</i> | 73 |

| | | |
|----------|---|------------|
| 4.2.4.3 | Resolución estructural y refinamiento | 74 |
| 4.2.5 | <i>Análisis estructural de SIGN-R1</i> | 75 |
| 4.2.5.1 | Descripción General de la Estructura | 75 |
| 4.2.5.2 | Sitio de unión a calcio | 76 |
| 4.2.5.3 | Comparación con otras lectinas homólogas de la familia – Ausencia de los calcios auxiliares | 77 |
| 4.2.5.4 | Presencia de moléculas de sulfato | 80 |
| 4.2.6 | <i>Descripción del sitio de unión a carbohidratos en SIGN-R1</i> | 81 |
| 4.2.6.1 | Complejo CRD_SIGN-R1 : Sulfodextrano..... | 81 |
| 4.2.6.2 | Complejo CRD_SIGN-R1 : Ácido siálico | 84 |
| 4.2.7 | <i>Cálculos de “docking” con ribitol fosfato y cápsula polisacárida</i> | 85 |
| 4.3 | DETERMINACIÓN ESTRUCTURAL DE LA PORCIÓN Fc..... | 86 |
| 4.3.1 | <i>Cristalización y optimización</i> | 86 |
| 4.3.2 | <i>Medida de los datos de difracción</i> | 87 |
| 4.3.3 | <i>Resolución estructural y refinamiento</i> | 88 |
| 4.3.4 | <i>Análisis de la estructura tridimensional</i> | 89 |
| 4.3.4.1 | Descripción general..... | 89 |
| 4.3.4.2 | Análisis estructural de las cadenas glicánicas | 90 |
| 5 | DISCUSIÓN | 95 |
| 5.1 | ESTRUCTURA DE LA ENDOLISINA Cpl-7 | 95 |
| 5.1.1 | <i>Producción de Cpl-7 como proteína de fusión con la lectinas SLS y cristalización</i> | 95 |
| 5.1.2 | <i>El módulo catalítico de Cpl-7 sigue el plegamiento de las GH-25</i> | 96 |
| 5.1.3 | <i>Reconocimiento del sustrato por el módulo catalítico de Cpl-7</i> | 96 |
| 5.1.4 | <i>Módulo de unión a la pared de Cpl-7</i> | 98 |
| 5.1.5 | <i>Especificidad de Cpl-7 por su sustrato</i> | 98 |
| 5.1.6 | <i>Disposición modular de la endolisina Cpl-7</i> | 99 |
| 5.2 | IMPLICACIONES BIOLÓGICAS DEL RECEPTOR SIGN-R1 | 101 |
| 5.2.1 | <i>SIGN-R1 en la inhibición del proceso inflamatorio</i> | 101 |
| 5.2.2 | <i>SIGN-R1 en la activación de la vía clásica del complemento</i> | 102 |
| 6 | CONCLUSIONES | 109 |
| 7 | BIBLIOGRAFÍA | 113 |
| 8 | APÉNDICE | 129 |

ABREVIATURAS

| | |
|---|--|
| ADN: Ácido desoxirribonucleico | IgG: Inmunoglobulina de tipo G |
| CBM: Módulo de unión a colina | LTA: Ácido lipoteicoico |
| CBPs: Proteínas de unión a colina | MC: Módulo catalítico |
| CBRs: Repeticiones de unión a colina | NeuAc: Ácido siálico |
| CPS: Cápsula polisacárida de <i>S. pneumoniae</i> | MurNAc: Ácido-N-acetilmurámico |
| C-t: Extremo C-terminal | N-t: Extremo N-terminal |
| CRD: Dominio de unión a carbohidratos | OD: Densidad óptica |
| CTLDs: Dominio lectinas de tipo C | PBS: Tampón salino fosfato |
| CWBM: Módulo de unión a la pared | PC: Fosforilcolina |
| DC: Células dendríticas | PDB: Base de datos de proteínas |
| DC-SIGN: <i>Dendritic cell specific intercellular adhesions molecules-3 (ICAM-3)-grabbing nonintegrin</i> | PG: Peptidoglicano |
| ECD: Dominio extracelular (extracellular domain) | r.m.s.d: Desviación cuadrática media |
| EDTA: Ácido etilendiaminetetraacético | SAXS: Small angle X-ray scattering |
| ESRF: European Synchrotron Radiation Facility | SDS-PAGE: Electroforesis en gel de poliacrilamida- |
| Fab: Región variable del anticuerpo | SLS: Lectina del hongo <i>Laetiporus sulphureus</i> |
| Fc: Región constante del anticuerpo | TA: Ácido teicoico |
| FPLC: Fast protein liquid chromatography | TEV: Virus jaspeado del tabaco |
| GlcNAc: N-acetilglucosamina | (2S5P)3: (GlcNAc-MurNAc-(L-Ala-D-isoGln-L-Lys-D-Ala-D-Ala))3 |
| GOLD: Genetic optimization for ligand docking | 2S2P: GlcNAc-MurNAc-L-Ala-D-isoGln |

INTRODUCCIÓN

1 INTRODUCCIÓN

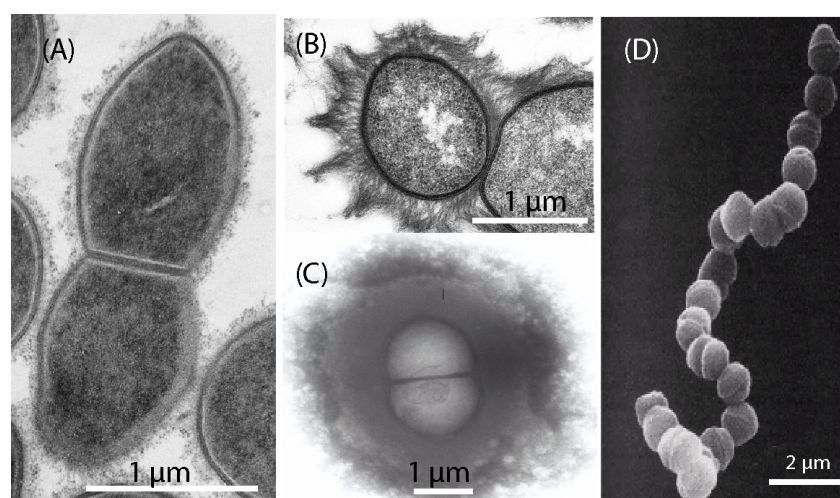
1.1 *Streptococcus pneumoniae*

1.1.1 Características generales

Streptococcus pneumoniae (neumococo) es una bacteria Gram positiva que posee una de las mayores tasas de patogenicidad para el ser humano. El Neumococo es el causante de procesos infecciosos, tales como neumonía, otitis media, sinusitis, peritonitis y de procesos invasivos severos, como meningitis y septicemia, afectando particularmente a niños, ancianos y personas inmunodeprimidas. La neumonía es, además, la mayor causa de mortalidad mundial en niños menores de 5 años, alcanzando un 29% del total de la tasa de mortalidad de esta fracción de la población, por encima de otras enfermedades como la malaria y el síndrome de inmunodeficiencia adquirida. El neumococo es el responsable del 10-25% de los casos de neumonía en países desarrollados, y de un 70% en países en vías de desarrollo (Tessa Wardlaw, 2006). Por todo ello, la Organización Mundial de la Salud ha establecido a partir del año 2008, un plan de acción mundial para la prevención y el control de la neumonía.

El neumococo es un diplococo de aproximadamente 1.2 – 1.8 μm de longitud, que tiene como hábitat natural la nasofaringe humana, en donde puede pasar un periodo determinado de tiempo sin ser infectivo, pero en condiciones adecuadas, el neumococo es capaz de iniciar procesos invasivos (Cartwright, 2002). Esta bacteria presenta diversas morfologías, la variante rugosa, desposeída de cápsula y la variante “lisa” de Griffith, la más frecuentemente aislada en personas infectadas con neumococo (fue precisamente Griffith en 1928 quien puso de manifiesto que la cápsula es el principal factor de virulencia). Además, existe una variedad capaz de producir filamentos de gran longitud (Dawson, 1933) (figura 1.1).

Figura 1.1. Morfología de neumococo. (A) Micrografía electrónica de neumococo estirpe rugosa (sin cápsula). (B) y (C) Micrografía electrónica de neumococo, estirpe lisa (encapsulado), serotipo 6B y 19F, respectivamente. (D) neumococo de larga cadena, filamento. <http://www.ppdictionary.com/bacteria/gpbac/pneumoniae.htm>.



La relevancia adquirida por las infecciones causadas por neumococo, unido a la inexistencia de una vacuna realmente eficaz o a la aparición de cepas resistentes a múltiples antibióticos, hacen necesario el desarrollo de nuevas estrategias terapéuticas. Así, recientemente ha cobrado especial

importancia el estudio de las proteínas de superficie de neumococo por su posible implicación en la virulencia de la bacteria y su potencialidad como agentes antigénicos (Hermoso *et al.*, 2007), así como también el estudio de receptores específicos implicados en la respuesta inmunitaria (Kang *et al.*, 2006). Otra de las estrategias cuya aplicación ha sido demostrada en estudios *in vivo* consiste en el empleo de proteínas líticas codificadas por los bacteriófagos que infectan a neumococo, como los recientemente introducidos enzibióticos (Pérez-Dorado *et al.*, 2007).

Una de las características más importantes del neumococo es su requerimiento nutricional absoluto por colina. La colina es incorporada en forma de fosforilcolina (PC) y aparece en los ácidos teicoicos (TAs) de la pared celular y en los ácidos lipoteicoicos (LTAs) de la membrana (Tomasz, 1967). Aunque en un principio se pensó que la virulencia del neumococo venía exclusivamente determinada por la presencia de la cápsula (Austrian, 1953), recientemente, se han descrito nuevos factores de virulencia, entre los que se encuentran diversas proteínas y componentes de la pared celular (McCullers & Tuomanen, 2001; Mitchell, 2003).

A continuación se describen detalladamente los distintos componentes estructurales de la envoltura celular en neumococo.

1.1.2 La cápsula polisacárida

La cápsula polisacárida de *S. pneumoniae* (CPS) es la estructura más externa y el principal factor de virulencia del neumococo. Hacia el interior le siguen la pared bacteriana, la membrana citoplasmática y el citoplasma, en donde está contenido el material genético (figura 1.2). La cápsula está compuesta principalmente por polisacáridos cargados negativamente a pH fisiológico, que rodean a la célula con un espesor variable entre 200 y 400 nm (Wartha *et al.*, 2007). La cápsula se mantiene en la superficie de la bacteria mediante enlaces covalentes a una ligasa que está adherida a la membrana citoplasmática (Kadioglu *et al.*, 2008), (figura 1.2).

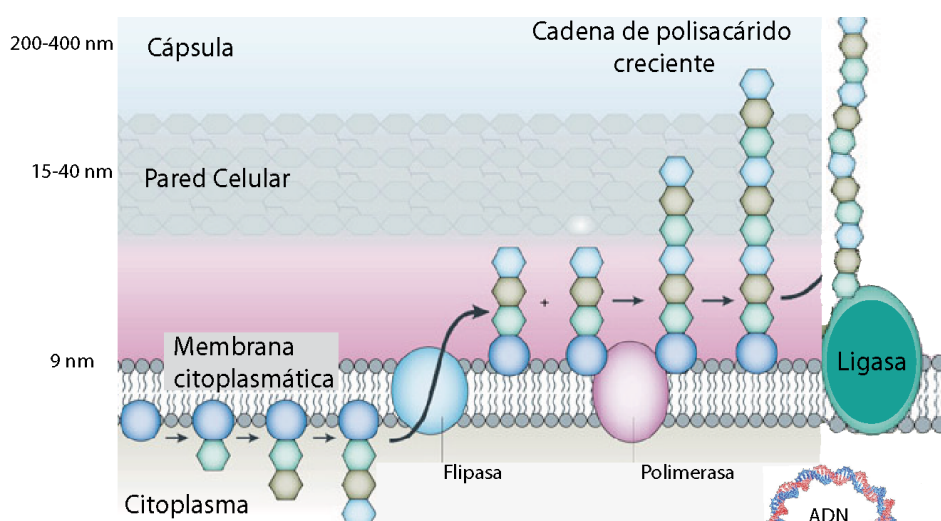


Figura 1.2. Esquema de la estructura externa de neumococo, donde se resume la síntesis de la cápsula polisacárida. Se describen las capas desde el exterior de la bacteria, cápsula, pared celular, membrana plasmática y citoplasma. Los espesores aproximados de cada capa se indican a la izquierda de la figura. La síntesis de la cápsula polisacárida comienza en el citoplasma en donde se sintetizan fragmentos pequeños, luego son translocados hacia el espacio exterior por la flipasa y allí son polimerizados y finalmente unidos a la ligasa que mantiene las cadenas de polisacárido en la superficie.

Se han descrito aproximadamente 90 serotipos capsulares distintos, con una composición química compleja y variable. En la figura 1.3 se listan algunos de los serotipos de los cuales se conoce su estructura (Jones, 2005). Los serotipos capsulares 1, 19A, 3, 5, 19F, 6 y 14 aparecen con mayor frecuencia en pacientes con neumonía y se consideran, por tanto, las variantes más virulentas (Resti *et al.*, 2010). La virulencia del neumococo está directamente relacionada con el grosor de la cápsula, característico de cada serotipo (Mac & Kraus, 1950), y con el aumento del grosor también aumenta la capacidad para bloquear el reconocimiento por parte del hospedador, impidiendo así la fagocitosis y la efectividad de la respuesta inmune (Jonsson *et al.*, 1985).

| Serotipo | Estructura de la cápsula polisacárida |
|--|--|
| <i>S. pneumoniae</i> Type 1 Stroop et al. 2002 | $\rightarrow 3\text{-D-AAT-}\alpha\text{-Galp-(1}\rightarrow 4\text{)-}\alpha\text{-D-GalpA(2/3OAc)-(1}\rightarrow 3\text{)-}\alpha\text{-D-GalpA-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 2 Jansson et al. 1998 | $\rightarrow 4\text{-}\beta\text{-D-Glcp-(1}\rightarrow 3\text{)-}[\alpha\text{-D-GlcpA-(1}\rightarrow 6\text{)-}\alpha\text{-D-Glcp-(1}\rightarrow 2\text{)]}$ $\alpha\text{-L-Rhap-(1}\rightarrow 3\text{)-}\alpha\text{-L-Rhap-(1}\rightarrow 3\text{)-}\beta\text{-L-Rhap-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 3 Reeves and Goebel 1941 | $\rightarrow 3\text{-}\beta\text{-D-GlcA-(1}\rightarrow 4\text{)-}\beta\text{-D-Glcp-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 4 Jones et al. 1991 | $\rightarrow 3\text{-}\beta\text{-D-ManpNAc-(1}\rightarrow 3\text{)-}\alpha\text{-L-FucpNAc-(1}\rightarrow 3\text{)-}\alpha\text{-D-GalpNAc-(1}\rightarrow 4\text{)}$ $\alpha\text{-D-Galp2,3(S)Py-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 5 Jansson et al. 1985 | $\rightarrow 4\text{-}\beta\text{-D-Glcp-(1}\rightarrow 4\text{)-}[\alpha\text{-L-PnepNAc-(1}\rightarrow 2\text{)-}\beta\text{-D-GlcpA-(1}\rightarrow 3\text{)]}$ $\alpha\text{-L-FucpNAc-(1}\rightarrow 3\text{)-}\beta\text{-D-Sugp-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 6B Kenne et al. 1979 | $\rightarrow 2\text{-}\alpha\text{-D-Galp-(1}\rightarrow 3\text{)-}\alpha\text{-D-Glcp-(1}\rightarrow 3\text{)-}\alpha\text{-L-Rhap-(1}\rightarrow 4\text{)-D-Rib-ol-(5}\rightarrow \text{P}\rightarrow$ |
| <i>S. pneumoniae</i> Type 9N Jones et al. 1985 | $\rightarrow 4\text{-}\alpha\text{-D-GlcpA-(1}\rightarrow 3\text{)-}\alpha\text{-D-Glcp-(1}\rightarrow 3\text{)-}\beta\text{-D-ManpNAc-(1}\rightarrow 4\text{)}$ $\beta\text{-D-Glcp-(1}\rightarrow 4\text{)-}\alpha\text{-D-GlcpNAc-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 9V Rutherford et al. 1991 | $\rightarrow 4\text{-}\alpha\text{-D-GlcpA(2/3OAc)-(1}\rightarrow 3\text{)-}\alpha\text{-D-Galp-(1}\rightarrow 3\text{)-}\beta\text{-D-ManpNAc(4/6OAc)-(1}\rightarrow 4\text{)}$ $\beta\text{-D-Glcp-(1}\rightarrow 4\text{)-}\alpha\text{-D-Glcp-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 12F Leontin et al. 1981 | $\rightarrow 4\text{-}[\alpha\text{-D-Galp-(1}\rightarrow 3\text{)]}\alpha\text{-L-FucpNAc-(1}\rightarrow 3\text{)-}\beta\text{-D-GlcNAc-(1}\rightarrow 4\text{)-}[\alpha\text{-D-Glc-(1}\rightarrow 2\text{)}$ $\alpha\text{-D-Glc-(1}\rightarrow 3\text{)]}\beta\text{-D-ManNAcA-(}\rightarrow$ |
| <i>S. pneumoniae</i> Type 14 Lindberg et al. 1977 | $\rightarrow 4\text{-}\beta\text{-D-Glcp-(1}\rightarrow 6\text{)-}[\beta\text{-D-Galp-(1}\rightarrow 4\text{)]}\beta\text{-D-GlcpNAc-(1}\rightarrow 3\text{)-}\beta\text{-D-Galp-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 18C Lindberg 1990b | $\rightarrow 4\text{-}\beta\text{-D-Glcp-(1}\rightarrow 4\text{)-}[\alpha\text{-D-Glcp(6OAc) (1}\rightarrow 2\text{)]}[\text{Gro-(1}\rightarrow \text{P}\rightarrow 3\text{)]}\beta\text{-D-Galp-(1}\rightarrow 4\text{)}$ $\alpha\text{-D-Glcp-(1}\rightarrow 3\text{)-}\beta\text{-L-Rhap-(1}\rightarrow$ |
| <i>S. pneumoniae</i> Type 19F Jennings et al. 1980 | $\rightarrow 4\text{-}\beta\text{-D-ManpNAc-(1}\rightarrow 4\text{)-}\alpha\text{-D-Glcp-(1}\rightarrow 2\text{)-}\alpha\text{-L-Rhap-(1}\rightarrow \text{P}\rightarrow$ |
| <i>S. pneumoniae</i> Type 23F Richards and Perry 1988 | $\rightarrow 4\text{-}\beta\text{-D-Glcp-(1}\rightarrow 4\text{)-}[\alpha\text{-L-Rhap-(1}\rightarrow 2\text{)]}\text{Gro-(2}\rightarrow \text{P}\rightarrow 3\text{)]}$ $\beta\text{-D-Galp-(1}\rightarrow 4\text{)-}\beta\text{-L-Rhap-(1}\rightarrow$ |

Figura 1.3. Representación esquemática de la estructura de la cápsula polisacárida para serotipo descrito hasta el momento presentes en *S. pneumoniae*.

Dependiendo de la composición de la envoltura bacteriana de neumococo, éste exhibe una virulencia diferenciada. En sangre las cepas de mayor virulencia presentan una mayor cantidad de cápsula polisacárida y menor contenido de ácidos teicoicos en la pared celular (Kim & Weiser, 1998), mientras que en la nasofaringe se hallan cepas que poseen menor cantidad de cápsula y mayor contenido de ácidos teicoicos (Weiser *et al.*, 1994).

1.1.3 La pared celular

La pared celular rodea la membrana citoplasmática y confiere a la bacteria una morfología típica. La adquisición de este tipo de estructuras, a modo de exoesqueleto, ha servido a las bacterias como mecanismo de adaptación al medio externo, protegiéndola de su posible lisis, ya sea osmótica o mecánica (Salton, 1994). Además de ser el medio de intercambio de solutos entre el exterior y el interior celular, la pared bacteriana sirve de punto de unión para toda una batería de proteínas

implicadas en procesos de crecimiento y división celular. De hecho, la pared desempeña un papel importante en los procesos de colonización, adherencia e invasión (McCullers & Tuomanen, 2001).

La composición molecular de la pared varía de unas especies a otras. En el caso de neumococo está constituida por un entramado de cadenas de peptidoglicano (mureína) y los ácidos teicoicos asociados a ellas, formando una estructura multilaminar con un espesor comprendido entre 15-40 nm (figura 1.4).

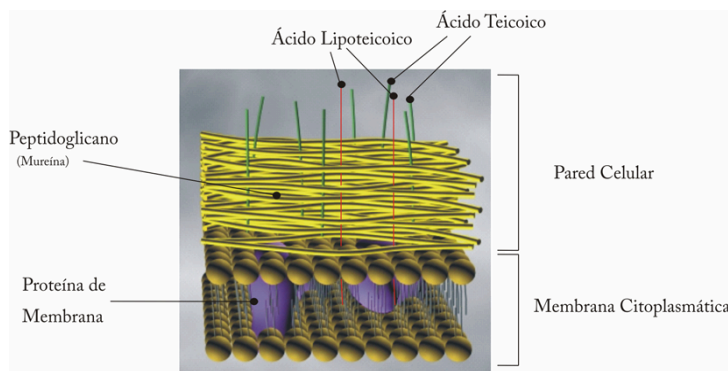


Figura 1.4. Representación esquemática de la pared celular de *S. pneumoniae*. Está constituida por el peptidoglicano (fibras amarillas) y los ácidos teicoicos (trazos verdes), ácidos lipoteicoicos (trazos rojos), así como la membrana citoplasmática (esferas marrones) y proteínas de membrana (morado).

1.1.3.1 El peptidoglicano

El peptidoglicano (PG) de neumococo está formado por un entramado tridimensional de cadenas glicánicas constituidas por residuos alternantes de ácido N-acetilmurámico (MurNAc) y glucosamina (GlcNAc), unidos mediante enlaces glicosídicos β -(1,4), que se entrecruzan mediante cortos segmentos peptídicos (figura 1.5). Cada cadena glicánica posee alrededor de unos 35 disacáridos, lo que supone, en su conformación más extendida, una longitud media de 35 nm (Beveridge, 1981). El nivel de acetilación de los residuos es del 90% para el murámico y del 16% para la glucosamina y su desacetilación tiene lugar tras su incorporación a la pared celular, lo cual constituye uno de los mecanismos de virulencia de la bacteria, al hacerla resistente a las lisozimas del hospedador (Vollmer & Tomasz, 2000).

En neumococo las cadenas peptídicas contienen mayoritariamente tres (L-Ala-D-iGln-L-Lys) o cuatro (L-Ala-D-iGln-L-Lys-D-Ala) aminoácidos, apareciendo en mucha menor proporción el pentapéptido L-Ala-D-iGln-L-Lys-D-Ala-D-Ala (Schleifer & Kandler, 1972). Existen dos tipos de conexiones interpeptídicas: la primera tiene lugar mediante el puente L-Ala-L-Ser (o en algunos caso L-Ala-L-Ala) que une la D-Ala en posición terminal de la unidad donadora y la L-Lys en posición 3 de la unidad receptora; el segundo tipo es el enlace directo entre la D-Ala terminal y la L-Lys (figura 1.5-A). Aunque estos aminoácidos son los que se han encontrado en mayor proporción en las uniones interpeptídicas, se ha descrito también la presencia minoritaria de glicina y aspártico (Fischer & Tomasz, 1985). Por otro lado, se ha observado que las estirpes resistentes a penicilina presentan un contenido importante de péptidos ramificados, mientras que en cepas sensibles el 70% de los péptidos son lineales (Severin & Tomasz, 1996).

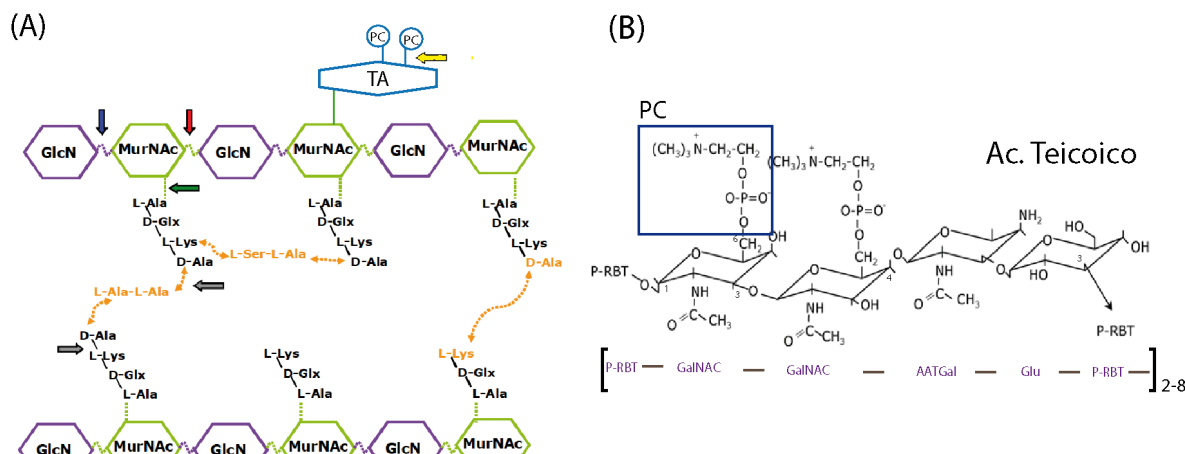
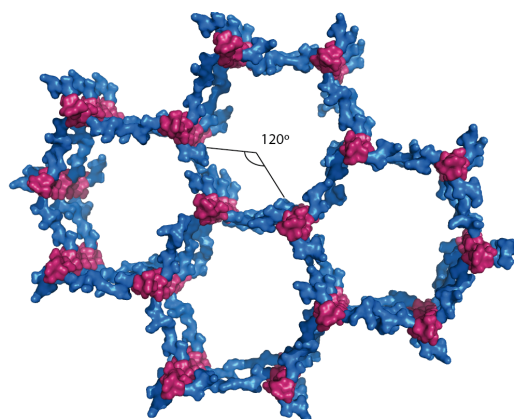


Figura 1.5. Esquema de la pared celular de neumococo. (A) Cadenas glicánicas formadas por repeticiones de glucosamina (GlcN) y ácido N-acetilmurámico (MurNAc) que se entrecruzan a través de segmentos peptídicos (negro) unidos a los residuos de MurNAc; las conexiones interpeptídicas (amarillo) que pueden ser directas o por puente peptídico. Unidos al MurNAc están los ácidos teicoicos (TA) que contienen fosforilcolina (PC). Las flechas indican los enlaces de la pared que son hidrolizados por distintas murein-hidrolasas: glucosaminidasas (azul), muramidasa o lisozimas (rojo), amidasa (verde), peptidasas (gris) y fosforilcolín-esterasa (amarillo). (B) Estructura primaria de una cadena de ácido teicoico (GalNAC, N-acetilgalactosamina; AATGal, diacetamidotridesoxigalactosa; Glu, glucosa; PC, fosforilcolina; P-RBT, ribitol fosfato).

La unión entre el polisacárido y las pequeñas cadenas peptídicas se realiza mediante el grupo D-láctico del ácido murámico. La disposición helicoidal que adoptan las cadenas glicánicas hace que las unidades peptídicas queden proyectadas en las tres direcciones del espacio, lo que permite establecer enlaces cruzados con los péptidos de cadenas próximas. Recientemente se ha determinado la estructura tridimensional de un fragmento sintético del peptidoglicano (Meroueh *et al.*, 2006), a partir del cual se propone que las cadenas glicánicas se dispondrían perpendicularmente a la membrana, con las cadenas peptídicas giradas cada 120°, permitiendo así la interacción con los péptidos adyacentes y formando en conjunto una estructura tipo panal de abeja (figura 1.6).

Figura 1.6. Estructura tridimensional de un fragmento sintético de peptidoglicano. Cadenas glicánicas (rosa) y las uniones peptídicas (azul).



1.1.3.2 Ácido teicoico y otros componentes de la pared celular

La envuelta celular de las bacterias Gram positivas contiene TAs y LTAs, polímeros con estructuras muy diversas, generalmente específicas de la especie, que le confieren características diferenciales a estas bacterias (Weidenmaier & Peschel, 2008). Los TAs de neumococo fueron descritos por primera vez en 1930 bajo la denominación de polisacárido C, debido a su interacción con la proteína C reactiva CRP (de sus siglas en inglés: *C reactive protein*) humana y están unidos a la pared celular. Los LTAs, se encuentran anclados a la membrana celular y, a diferencia de lo que

ocurre con otras bacterias Gram positivas, neumococo es un caso único ya que los LTAs tienen la misma composición que los TAs (Fischer, 1993).

Los TAs de neumococo son polímeros de naturaleza zwitteriónica cuya unidad repetitiva, según el modelo hasta ahora aceptado, es el pentasacárido formado por D-glucosa (Glc), D-2-acetamido-4-amino-2,4,6-tridesoxigalactosa (AATGal), D-N-acetilglucosamina (GalNAc), D-N-acetilglucosamina y D-ribitol-5-fosfato (RBT), y los enlaces glicosídicos entre los distintos azúcares son, ($-\beta$ -1,3), ($-\alpha$ -1,4), ($-\alpha$ -1,3) y ($-\beta$ -1,1) respectivamente, esquematizándose su estructura en la figura 1.5-B. Las unidades de TA se conectan entre sí a través del residuo de ribitol, el cual se une a la posición 6 del anillo de glucosa de la unidad anterior a través un enlace fosfodiéster (Fischer *et al.*, 1993) y cada cadena de teicoico contiene entre 6 y 8 repeticiones del pentasacárido antes descrito (Yother *et al.*, 1998). La presencia de los grupos fosfato confiere a los TAs y LTAs la capacidad de formar quelatos con magnesio y otros cationes divalentes, lo que contribuye a aumentar la rigidez del exoesqueleto. Como en muchas otras bacterias Gram positivas, los TAs de neumococo también contienen D-Ala (Draing *et al.*, 2006) y esta modificación confiere resistencia a la bacteria frente a la acción de péptidos antimicrobianos de naturaleza catiónica.

La unión del TA al peptidoglicano se establece a través de un conector formado por glicerolfosfato, ribitolfosfato o glucosilfosfato que puede estar esterificado o glicosilado (su composición en neumococo no ha sido determinada aún). Este conector se une, a su vez, a la cadena glicánica a través de un disacárido que forma un enlace fosfodiéster con la posición 6 del MurNAc del peptidoglicano. La incorporación del TA al peptidoglicano tiene lugar durante la formación de la pared celular (Fischer & Tomasz, 1984) y el contenido medio estimado es de una cadena de TA con ocho repeticiones por cada tetradisacárido tetrapéptido (García-Bustos *et al.*, 1987). El LTA se inserta en la membrana plasmática uniéndose covalentemente a un glicolípido cuya estructura ha sido recientemente determinada (Seo *et al.*, 2008).

Los residuos de fosforilcolina presentes en TA y LTA, que participan directamente en la internalización de la bacteria (Cundell *et al.*, 1995), son reconocidos durante el proceso de inflamación por la CRP humana (Mold *et al.*, 1982) y actúan como receptores para toda una familia de proteínas codificadas por neumococo, denominadas proteínas de unión a colina (CBPs, de sus siglas en inglés: *choline binding proteins*), entre las que se encuentran las enzimas que hidrolizan la pared celular (mureín-hidrolasas) y otros factores implicados en la invasión de la bacteria (Jedrzejewski, 2001). Tanto los TAs como los LTAs son altamente antigénicos, activan la ruta alternativa del complemento (Hummell *et al.*, 1985), estimulan la formación de interleucina-1 en los monocitos humanos (Riesenfeld-Orn *et al.*, 1989) y actúan como agentes quimiotácticos en la inflamación meníngea (Tuomanen *et al.*, 1985). Forman parte, además, de los receptores de algunos bacteriófagos (López *et al.*, 1982).

1.1.4 Proteínas de superficie en *S. pneumoniae*

A la espera de un análisis de proteómica más detallado en *S. pneumoniae*, su secuencia genómica reveló la existencia de tres grupos principales de proteínas de superficie: la familia de las proteínas de unión a colina (CBPs), las lipoproteínas y las proteínas con motivos aminoacídicos LPXTG que están covalentemente unidas a la pared celular. Recientemente se han descrito también las proteínas de superficie no clásicas, de las cuales se desconocen los motivos de unión a superficie (Bergmann & Hammerschmidt, 2006) (figura 1.7). Muchas de estas proteínas de superficie son factores de virulencia que contribuyen a la patogénesis de este organismo. Además, estas proteínas participan en interacciones específicas con los tejidos del hospedador, facilitando la supervivencia bacteriana, ayudando a la expansión de la bacteria dentro de los tejidos del hospedador, y ocultando la superficie bacteriana de los mecanismos de defensa del hospedador. Debido a estas propiedades, numerosas proteínas expuestas a la superficie son dianas potenciales para el diseño de agentes profilácticos tales como vacunas y algunos de ellos podrían servir como dianas para el diseño de drogas terapéuticas.

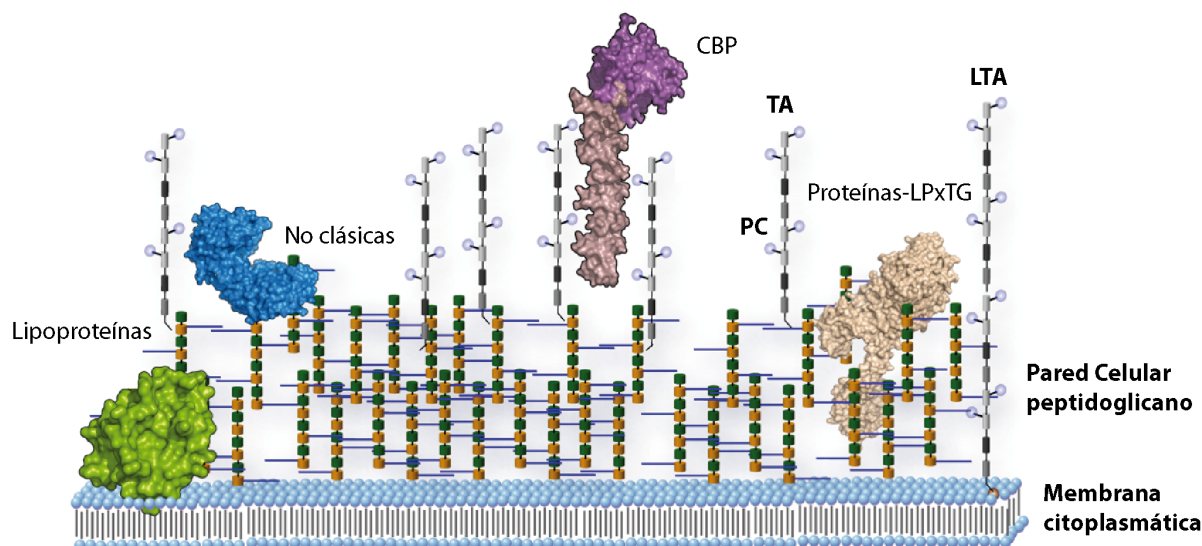


Figura 1.7. Modelo esquemático de la pared celular de neumococo y sus proteínas de superficie. Unidas a la membrana plasmática, se anclan las lipoproteínas. En el peptidoglicano se anclan covalentemente las proteínas con motivos LPXTG. En los ácidos teicoicos y lipoteicoicos, los cuales contienen la fosforilcolina (PC), se anclan de forma no covalente las proteínas de unión a colina (CBP). También se señala el grupo de proteínas no clásicas de las cuales se desconocen los motivos de reconocimiento.

1.1.5 Enzimas líticas de pared de neumococo y sus fagos: Cpl-7 un caso especial

S. pneumoniae sintetiza una serie de enzimas dependientes de colina implicadas en la hidrólisis de determinados enlaces del peptidoglicano (Giesbrecht *et al.*, 1976), llamadas mureín-hidrolasas. Éstas participan en el crecimiento celular, la renovación de la pared y la separación de las células hijas (García-Bustos & Tomasz, 1987). La actividad lítica de las CBPs está finamente regulada a lo largo del ciclo celular y se encuentran frecuentemente asociadas a la membrana o interaccionando fuertemente con la envuelta celular (Cleveland *et al.*, 1975; Holtje & Tuomanen, 1991).

Todos los bacteriófagos de doble cadena codifican, al menos, para una enzima lítica (endolisina) utilizada para liberar y dispersar la progenie viral al final del ciclo lítico (Borysowski *et al.*, 2005). La actividad de esta endolisina, está regulada únicamente por un gen tardío que codifica para una pequeña proteína denominada holina, que facilita el acceso de la enzima a la pared de la bacteria mediante la creación de poros en la membrana (Young, 1992). Las lisinas codificadas por neumococo y sus bacteriófagos son proteínas modulares relacionadas estructural y funcionalmente, habiéndose postulado un mecanismo de intercambio modular a lo largo de la evolución entre los genes que las codifican (Romero *et al.*, 1990; Díaz *et al.*, 1991; Patthy, 1991). Como se muestra en la figura 1.8, todas contienen un módulo catalítico y otro módulo de unión a la pared (CWBM, de sus siglas en inglés: *cell wall binding module*) que determina, en gran medida, la especificidad hacia el sustrato. Esta organización modular incrementa la actividad enzimática sobre paredes intactas, haciéndola superior en varios órdenes de magnitud a las de los módulos catalíticos aislados (Sánchez-Puelles *et al.*, 1990; Sanz *et al.*, 1992a; García *et al.*, 1999). Basándose en el enlace hidrolizado, las mureín-hidrolasas se clasifican en:

Amidasas: Rompen el enlace amida entre el grupo D-lactilo del MurNAc y el grupo amino de la L-Ala. A este grupo pertenecen LytA, la autolisina mayoritaria de neumococo, y las endolisinas Hbl, Ejl y Pal, codificadas por los fagos Hb-3, Ej-1 y Dp-1, respectivamente.

Glucosaminidasas: Son glicosilhidrolasas que hidrolizan el enlace GlcNAc- β -(1,4)-MurNAc. A este grupo pertenece LytB, codificada por neumococo.

Transglicosidasas: Tienen el mismo tipo de actividad que las lisozimas pero generan 1,6-anhidromurámico, bloqueando el extremo reductor de la cadena glicánica. Hasta el momento, no se ha caracterizado ninguna transglicosidasa lítica codificada por neumococo.

Endopeptidasas y carboxipeptidasas: Hidrolizan los enlaces peptídicos, aunque las carboxipeptidasas se limitan a separar solamente el aminoácido C-terminal. Hasta la fecha, tampoco se ha identificado ninguna peptidasa específica de la pared celular de neumococo.

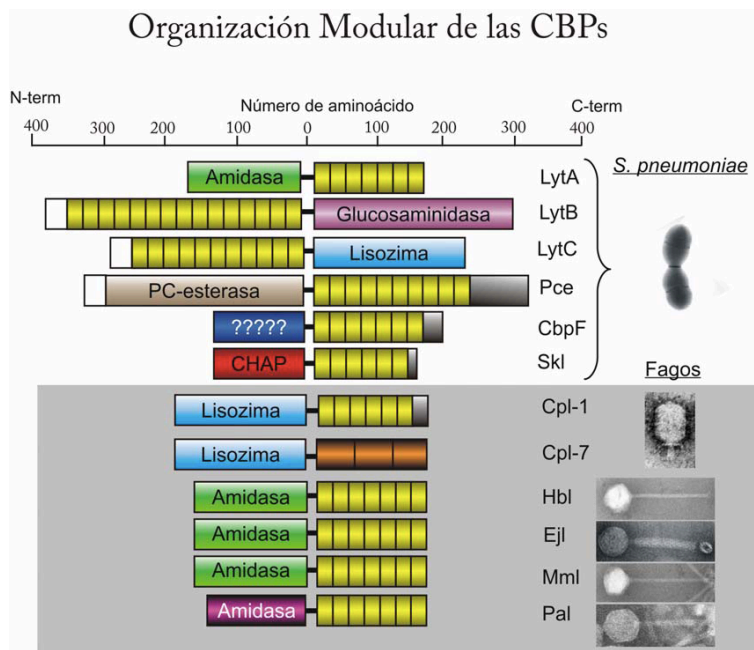
Fosforilcolín-esterasas: Dentro de la familia de las mureín-hidrolasas también se incluye la fosforilcolín-esterasa de neumococo (Pce) que libera residuos de fosforilcolina de los TAs y LTAs.

Muramidasas o lisozimas: También pertenecen a la familia de las glicosilhidrolasas. Son β -N-acetilmuramidasas y rompen el enlace glicosídico MurNAc- β -(1,4)GlcNAc. Dentro de este grupo se encuentran LytC, la lisozima codificada por neumococo, y las endolisinas Cpl-1, Cpl-7 y Cpl-9, codificadas por los bacteriófagos de neumococo Cp-1, Cp-7 y Cp-9, respectivamente.

Todas las CBPs y todas las endolisinas de fagos de neumococo, a excepción de Cpl-7, presentan un módulo de unión a la pared que une de forma específica los residuos de fosforilcolina de los TAs y LTAs, por lo que se le denomina módulo de unión a colina (CBM, de sus siglas en inglés: *choline binding module*). El CBM está constituido por la repetición en tándem de un motivo conservado de secuencia de unos 17-23 aminoácidos y podría considerarse como una unidad

modular surgida por duplicación genética de una secuencia original. Como puede verse en la figura 1.8, el número de repeticiones del CBM, su posición en la molécula y el grado de conservación de la secuencia, varía de unas enzimas a otras.

Figura 1.8. Esquema de la organización modular de las proteínas de unión a colina. Los dominios funcionales se identifican según su función; en verde (amidasa), rojo (amidasa de la familia CHAP), morado (glucosaminidasa), azul claro (lisozimas), azul oscuro (regulador) y marrón (fosforilcolina esterasa). Las repeticiones de los CBMs se muestran en amarillo y Cpl-7 con repeticiones que no unen colina en naranja.



Cpl-7 es un caso particular y único hasta la fecha dentro de la familia. Portadora de un módulo de unión a la pared que no tiene homología con los módulos de unión a colina, es capaz de hidrolizar indistintamente paredes que contengan fosforilcolina o etanolamina (García *et al.*, 1988; García *et al.*, 1990; Díaz *et al.*, 1991), lo que demuestra que su actividad no depende de la unión a colina. Estructuralmente, el módulo catalítico se encuentra en posición N-terminal y comparte un 85,6% de la secuencia con el de Cpl-1; el módulo de unión a la pared, en posición C-terminal, está formado por un motivo de 48 aminoácidos, que se repite de forma idéntica 2,8 veces y es completamente diferente a los módulos de unión a colina, y su especificidad de unión es desconocida hasta ahora.

Es de suma importancia destacar la identificación de proteínas implicadas en la hidrólisis de la pared celular bacteriana cuyas secuencias incluyen las repeticiones del tipo de Cpl-7. Este hecho sugiere que Cpl-7, podría ser, al igual que la amidasa Pal (Sheehan *et al.*, 1997), una quimera natural de origen intergenético formada por la fusión de dos genes: uno (presente en neumococo) codificaría para el módulo catalítico y el otro (procedente de un ancestro todavía desconocido) codificaría para la región implicada en el reconocimiento y unión a la pared celular, en donde Cpl-7 sería capaz de lisar otras bacterias. Por tanto, el estudio detallado de Cpl-7 podría aportar información relevante sobre su función y sentar las bases del reconocimiento a la pared celular, no sólo en neumococo sino también en otras bacterias patógenas.

1.1.5.1 Mecanismo de hidrólisis de las lisozimas de la familia GH-25

Las lisozimas o muramidasa se clasifican en cuatro familias (GH-22 a GH-25) en función de la similitud de secuencia y/o de la homología en su plegamiento tridimensional

(www.cazy.org/fam/acc_GH.html). Las dos primeras (GH-22 y GH-23) incluyen, respectivamente, a las enzimas relacionadas con las lisozimas de huevo de gallina (HEWL) y huevo de ganso (GEWL), mientras que las dos últimas (GH-24 y GH-25) tienen como prototipos a las lisozimas del fago T4 (T4L) y del hongo *Chalaropsis*. A pesar de que la similitud de secuencia entre las enzimas de las tres primeras familias es baja (Monzingo *et al.*, 1996; Robertus *et al.*, 1998), la topología de la cavidad catalítica es similar en todas ellas, habiéndose postulado un origen evolutivo común. La familia GH-25, en la que están incluidas las lisozimas de neumococo y sus bacteriófagos (Felch *et al.*, 1975; Lichenstein *et al.*, 1990), presenta, sin embargo, un plegamiento completamente distinto al de las otras tres familias (Hermoso *et al.*, 2003; Pérez-Dorado *et al.*, 2010).

Las familias GH-22, 23 y 24 también muestran modos comunes de unión al sustrato. Su cavidad catalítica puede acomodar hasta seis unidades de azúcar cuyas posiciones han sido denominadas de A a F o de -4 a +2, comenzando por el extremo no reductor de la cadena glicánica, y las posiciones -1 y +1 (subsitos D y E) son ocupadas por los anillos implicados en el enlace glicosídico a hidrolizar (Hadfield *et al.*, 1994). El muropéptido A (-4) B (-3) C (-2) D (-1) E (+1) F (+2) (figura 1.9), se une a lo largo de la cavidad adoptando una conformación distorsionada (Vocadlo *et al.*, 2001) en la que la unidad de MurNAc en -1 se encuentra girada respecto al anillo en posición +1, dando lugar a un cambio de dirección en la cadena que parece facilitar la catálisis (Monzingo *et al.*, 1996). A pesar de las similitudes, las tres familias muestran diferencias importantes tanto en la especificidad hacia el sustrato como en la actividad catalítica.

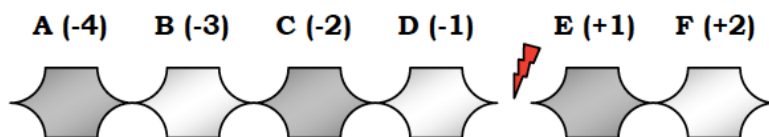


Figura 1.9. Esquema de los residuos de azúcar reconocidos en el sitio catalítico de HEWL. En la parte superior se indica el nombre (A-F) y la posición (-4 a +2) del subsito correspondiente en la cavidad catalítica, comenzando a partir del extremo no reductor. La flecha roja indica el enlace que va a ser hidrolizado.

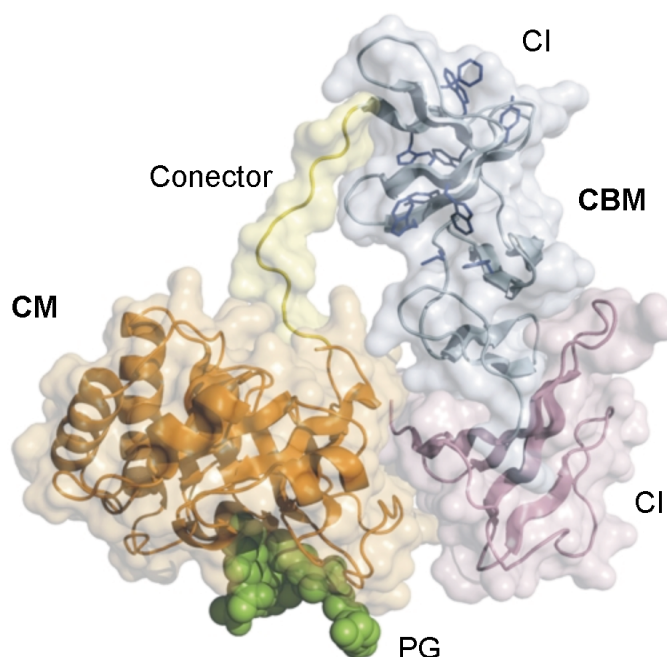
La unión del sustrato a las lisozimas de la familia GH-25 sólo se ha caracterizado estructuralmente en el caso de la endolisina Cpl-1, codificada por el bacteriófago Cp1 (Pérez-Dorado *et al.*, 2007) y la autolisina LytC (Pérez-Dorado *et al.*, 2010).

La estructura completa de la lisozima Cpl-1 ilustra claramente la naturaleza modular característica de las proteínas líticas de superficie de neumococo (Hermoso *et al.*, 2003) (figura 1.10). La cadena polipeptídica está formada por el módulo catalítico y el módulo de unión a colina, unidos mediante un conector ácido constituido por 10 residuos. El módulo catalítico está formado por un único dominio estructural que se pliega formando un barril α/β de estructura irregular (característico de la familia GH-25 de las glicosilhidrolasas) al que le faltan las tres últimas hélices α y en el que la hebra $\beta 8$ adopta una disposición antiparalela. Las repeticiones del CBM se organizan en dos regiones estructurales bien definidas, denominadas CI y CII (figura 1.10). El dominio CI está implicado en el reconocimiento de la colina en la pared y el dominio CII está directamente implicado en las interacciones que estabilizan los módulos que componen a Cpl-1.

La posición relativa de ambos módulos viene determinada por la estructura extendida del conector que los une, constituido mayoritariamente por residuos ácidos, y por las fuertes interacciones hidrofobias que se establecen entre el módulo catalítico y la región C-terminal del CBM (Hermoso *et al.*, 2003).

A pesar de que todas las repeticiones del CBM de Cpl-1 presentan el mismo tipo de estructura secundaria y de que los residuos implicados en la unión a colina están absolutamente conservados, la desviación del plegamiento global de CII con respecto a la superhélice a izquierdas hace que sólo dos de los cinco posibles sitios de unión a colina de Cpl-1 sean funcionales (Hermoso *et al.*, 2003). Este hecho pone de manifiesto la versatilidad de las repeticiones que forman los CBMs y su capacidad para actuar como receptores de colina o para mediar en las interacciones entre dominios.

Figura 1.10. Estructura tridimensional de Cpl-1. CM, módulo catalítico en naranja; conector en amarillo y CBM, módulo de unión a colina constituido por CI en gris y CII morado; PG, fragmento de peptidoglicano (PG) esferas verdes. (Pérez-Dorado *et al.*, 2007).



La interacción del módulo catalítico de Cpl-1 con sus sustratos ha sido caracterizada (Hermoso *et al.*, 2003), habiéndose se han identificado los residuos implicados en la catálisis, el Glu94 y el Asp10, los cuales son comunes para todos los miembros de la familia GH-25. La cavidad catalítica de Cpl-1 (figura 1.11) es capaz de acomodar, al menos, cinco unidades de azúcar (subsitos -2 a +3) en dos sitios de unión al PG. El sitio 1 de unión al PG reconocería a los anillos situados en las posiciones +1 a +3 y la cadena peptídica del MurNAc en +2. El sitio 2 de unión al PG reconocería los últimos anillos en posición -1 y -2 y la cadena peptídica del MurNAc en -1. El anillo de azúcar en posición -1 adopta una conformación en bote que introduce un giro en la cadena glicánica equivalente al descrito para los otros tres tipos de lisozimas (Pérez-Dorado *et al.*, 2007). A diferencia de lo que ocurre en éstas, el grupo acetamido del anillo de GlcNAc en posición -2 no es un elemento clave en la unión del sustrato, estando situado, de hecho, fuera de la cavidad catalítica (Hermoso *et al.*, 2003; Pérez-Dorado *et al.*, 2007). Este comportamiento explica la capacidad de las lisozimas de la familia GH-25 para hidrolizar paredes de neumococo, en las que el porcentaje de residuos de glucosamina

acetilados es del 10%, así como resistencia de la bacteria a las lisozimas de las otras tres familias, incluidas las enzimas del hospedador (Vollmer & Tomasz, 2000).

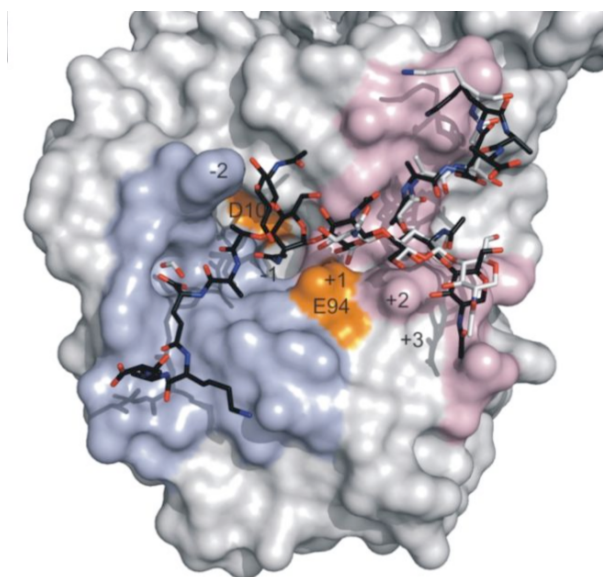


Figura 1.11. Detalles del reconocimiento del PG por Cpl-1. Las regiones involucradas en el reconocimiento del sustrato están coloreadas en rosa claro para el sitio de unión 1 y azul para el sitio de unión 2. Los residuos catalíticos Glu94 y Asp10 se muestran en naranja. Se destacan los subsitios de los azúcares así como también la zona de reconocimiento de la parte peptídica (Pérez-Dorado *et al.*, 2007).

La estructura completa de la endolisina LytC es la primera lisozima codificada por neumococo de la cual se ha reportado la estructura tridimensional en complejo con un fragmento de peptidoglicano purificado de neumococo (cepa R6) y colina (Pérez-Dorado *et al.*, 2010). LytC exhibe una estructura modular característica de las CBPs y en la figura 1.12-A se muestra la estructura completa que consta de un módulo catalítico y un módulo de unión a colina, el cual contiene nueve repeticiones, características y no canónicas, capaces de unir hasta 8 moléculas de colina. A diferencia de Cpl-1, en donde la conexión entre ambos módulos es a través de un bucle simple de 10 aminoácidos, en LytC esta conexión está constituida por dos repeticiones de unión a colina de 24 residuos (p10) y 25 residuos (p11), con estructura secundaria bien establecida. Este plegamiento único en LytC, obliga al sitio catalítico a quedar anclado al CBM, quedando así la cavidad catalítica poco expuesta al sustrato (figura 1.12-A). Esta característica estructural sugiere un mecanismo de activación y regulación *in vivo* único que depende de la activación de otras enzimas líticas y de superficie (Pérez-Dorado *et al.*, 2010).

LytC también pertenece a la familia GH-25 de las glicosilhidrolasas con lo que el mecanismo catalítico es similar a Cpl-1, en donde están involucrados dos residuos ácidos Glu365 y Asp273. Mientras que los residuos involucrados en la estabilización de la cadena glicánica están conservados, hay una notable diferencia en la estabilización de la porción peptídica unida al anillo de MurNAc en la posición -1, debido a que LytC posee un bucle (Lc) en esta zona, lo que distingue de Cpl-1 (figura 1.12-B) y que sugiere que los mecanismos de reconocimiento para el sustrato son diferentes en LytC y Cpl-1. Además el hecho de que el sitio activo de LytC esté enfrenteado al CBM, ello le proporciona un impedimento de su actividad lo que está estrechamente relacionado con su regulación (Pérez-Dorado *et al.*, 2010) (figura 1.12-C).

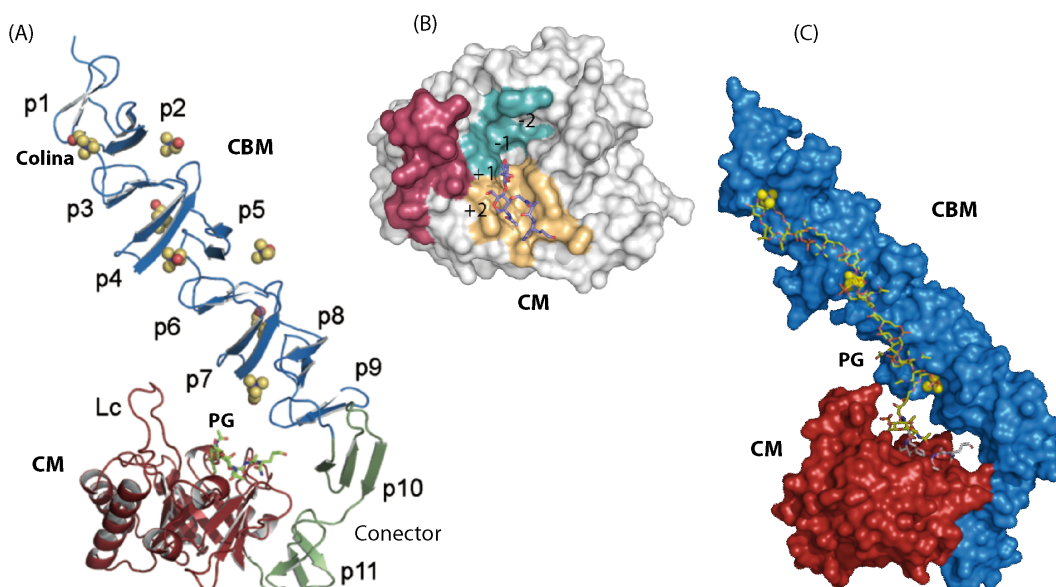


Figura 1.12. Estructura tridimensional de LytC en complejo con PG. (A) Estructura completa de LytC, el módulo catalítico (CM) en rojo; conector en verde (p10 y p11) y el módulo de unión a colina (CBM) con 9 repeticiones capaces de reconocer 8 moléculas de colina en amarillo; PG, fragmento de peptidoglicano, representado como varillas en verde. (B) Estructura de la cavidad catalítica de LytC, el sitio primario y secundario se muestran en amarillo y azul, y el nuevo bucle característico de LytC (Lc) en rojo. Se muestra el disacárido dipéptido GluN-MurNAc-Ala-glu en posición +1 y +2. (C) Modelo estructural del reconocimiento del peptidoglicano por ambos módulos de LytC. (Pérez-Dorado *et al.*, 2010).

Desde el punto de vista catalítico y tras la unión del sustrato, el anillo de la Tyr127 (Tyr407 en LytC), situado en el centro del canal que forma el sitio 1 de unión al PG, se desplaza para acomodar el sustrato (Pérez-Dorado *et al.*, 2007; Pérez-Dorado *et al.*, 2010) (figura 1.13).

En todas las lisozimas cuyo mecanismo catalítico ha sido estudiado la hidrólisis del enlace glicosídico transcurre mediante una reacción general ácido/base. Dependiendo de la geometría y la estereoquímica del sitio catalítico, y la reacción puede tener lugar con retención o con inversión de la configuración del carbono anomérico.

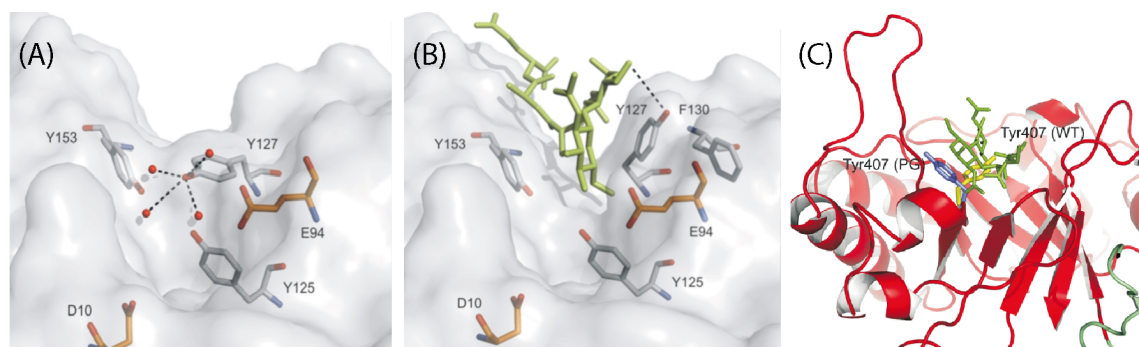


Figura 1.13. Estructura del canal de unión al sustrato en Cpl-1. (A) Estructura en ausencia de peptidoglicano Tyr127 bloquea la entrada del sustrato al canal. (B) En presencia del ligando se desplaza la Tyr127. El ligando y los residuos catalíticos se muestran en verde y naranja respectivamente, (Pérez-Dorado *et al.*, 2007). (C) Superposición de las estructuras en ausencia (estructura nativa, WT en color amarillo) y presencia de peptidoglicano (PG en azul) en LytC, Tyr407 (Pérez-Dorado *et al.*, 2010).

En ambos casos hay dos residuos catalíticos de carácter ácido que actúan uno como donador/aceptor de protones, y el otro como nucleófilo (retención de la configuración) o como base general (inversión de la configuración). La separación entre los residuos catalíticos, de 4.8 – 5.3 Å (para enzimas con retención) y 9.6 – 9.9 Å (para enzimas inversoras), determina el mecanismo de la reacción (Wang *et al.*, 1995). La retención de la configuración se debe a que la molécula de agua que

lleva a cabo el ataque nucleofílico actúa por la misma cara por la que se libera el grupo saliente (Tews *et al.*, 1996) (figura 1.14-A). En los dos mecanismos, el residuo que actúa como ácido general está situado a una distancia típica de enlace de hidrógeno del oxígeno glicosídico.

La reacción de hidrólisis en las lisozimas, como Cpl-1 y LytC, que actúan con inversión de la configuración, tiene lugar por un mecanismo de sustitución nucleofílica en una sola etapa (figura 1.14-B). En esta reacción, el residuo donador de protones se sitúa en la cara β del azúcar y protona al oxígeno del enlace glicosídico. Simultáneamente, la molécula de agua que lleva a cabo el ataque nucleofílico es activada por el residuo básico y reemplaza al grupo saliente, actuando por la cara α del anillo, (Weaver *et al.*, 1995; Kudva *et al.*, 1999).

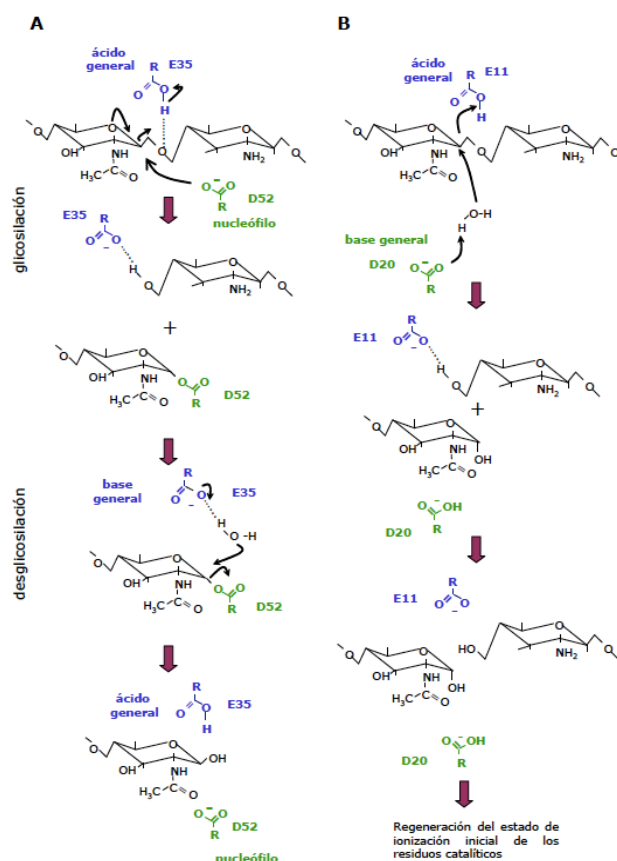


Figura 1.14. Mecanismos de reacción de las lisozimas. (A) Reacción con retención de configuración β del carbono anomérico (mecanismo Koshland). (B) Mecanismo con inversión de la configuración (β - α) del carbono anomérico. La numeración de los residuos corresponde a las lisozimas HEWL (A) y T4L (B).

1.1.6 Terapia fágica y enzibióticos

La capacidad de los antibióticos para curar enfermedades infecciosas, que en el pasado eran mortales, está viéndose seriamente amenazada en las últimas décadas debido al gran desarrollo de resistencias por parte de las bacterias que, en algunos casos, llegan a aparecer frente a varios antibióticos a la vez (multiresistencia). Por ello el desarrollo de tratamientos alternativos capaces de prevenir y erradicar las infecciones de origen bacteriano se ha convertido en una cuestión vital (Service, 2004). Entre las alternativas más prometedoras frente a las infecciones causadas por las bacterias Gram positivas se encuentran la terapia fágica y el uso de productos fágicos (endolisinas o “enzibióticos”). Una de las principales ventajas que éstos presentan es que son elementos

ampliamente distribuidos en la naturaleza y que poseen especificidad frente a las bacterias (Borysowski *et al.*, 2005; Fischetti, 2006; Fischetti *et al.*, 2006).

A principios del siglo XX se observó el efecto lítico de los bacteriófagos en las placas de cultivos, y al considerar su capacidad bactericida, se comenzó a pensar en ellos como un tratamiento frente a las infecciones en humanos y en animales (Twort & Twort, 1921; d'Herelle, 1931). A pesar de que durante los años 20 y 30 se intentó generalizar su uso, la terapia fágica cayó en el olvido (excepto en algunos países de Europa del Este) por diversas causas como el desconocimiento de la biología de los fagos, la falta de protocolos y controles adecuados y la industrialización de los antibióticos. En las últimas décadas ha resurgido el interés científico por este tipo de tratamiento debido al problema imparable de las resistencias bacterianas. Sin embargo, a pesar de que hay varios estudios que avalan el posible uso de la terapia fágica y sus ventajas (por ejemplo, en casos de bacterias resistentes a casi todo el arsenal de antibióticos (Markoishvili *et al.*, 2002)), sigue habiendo problemas que hay que tratar con prudencia: la dudosa eficacia de los fagos en infecciones donde éstos encuentren dificultades para su multiplicación (como en abscesos, en donde el número de células es excesivo, o en bacterias de multiplicación intracelular), el rápido desarrollo de resistencias frente a los fagos por parte de las bacterias (lo que implica utilizar cócteles de bacteriófagos), etc. (Kudva *et al.*, 1999; Fischetti *et al.*, 2006; Deresinski, 2009).

Una variante de la terapia fágica es la utilización como antibióticos de las endolisinas codificadas por los bacteriófagos que infectan a las bacterias Gram positivas como agentes antimicrobianos, conocidos desde entonces como enzibióticos. Esta nueva y prometedora alternativa supone un nuevo mecanismo de acción terapéutica (hidrólisis de la pared celular bacteriana) que incluye, también, a las autolisinas (Rodríguez-Cerrato *et al.*, 2007a) y que está siendo objeto de numerosos estudios (Hermoso *et al.*, 2007; Donovan & Foster-Frey, 2008). Al igual que en la terapia fágica, las endolisinas son altamente específicas por lo que una de las ventajas es que la microflora normal del hospedador no se vería afectada por su uso (Loeffler *et al.*, 2001).

Dada la versatilidad estructural de las endolisinas, éstas podrían recombinarse entre sí para crear agentes antimicrobianos con nuevas especificidades. Además, el desarrollo de resistencias frente a las enzimas líticas parece menos probable que en el caso de los fagos ya que esto implicaría remodelar elementos de la pared bacteriana que son esenciales para la célula (como la colina en el caso de neumococo) y, por tanto, la generación de estas resistencias sería inviable para la propia bacteria. Centrándonos en el caso de las endolisinas que degradan la pared de neumococo, se ha demostrado la capacidad de las endolisinas Cpl-1 y Pal y de la autolisina LytA para actuar como antimicrobianos eficaces y específicos en el tratamiento y prevención de las infecciones causadas por la bacteria (Loeffler *et al.*, 2001; Jado *et al.*, 2003; Loeffler *et al.*, 2003; Entenza *et al.*, 2005; Rodríguez-Cerrato *et al.*, 2007a), así como su capacidad para actuar sinérgicamente entre ellas y con antibióticos clásicos como la penicilina o la gentamicina (Djurkovic *et al.*, 2005; Rodríguez-Cerrato *et al.*, 2007b). Todas estas consideraciones hacen que la apuesta por los enzibióticos y su estudio cobre mayor interés cada día y ofrezca una posibilidad prometedora en la lucha contra las infecciones

bacterianas, y más aún si apuntamos a endolisinas con módulos de anclaje universales como es el caso de Cpl-7, el cual está presente en otras bacterias y bacteriófagos (Donovan & Foster-Frey, 2008).

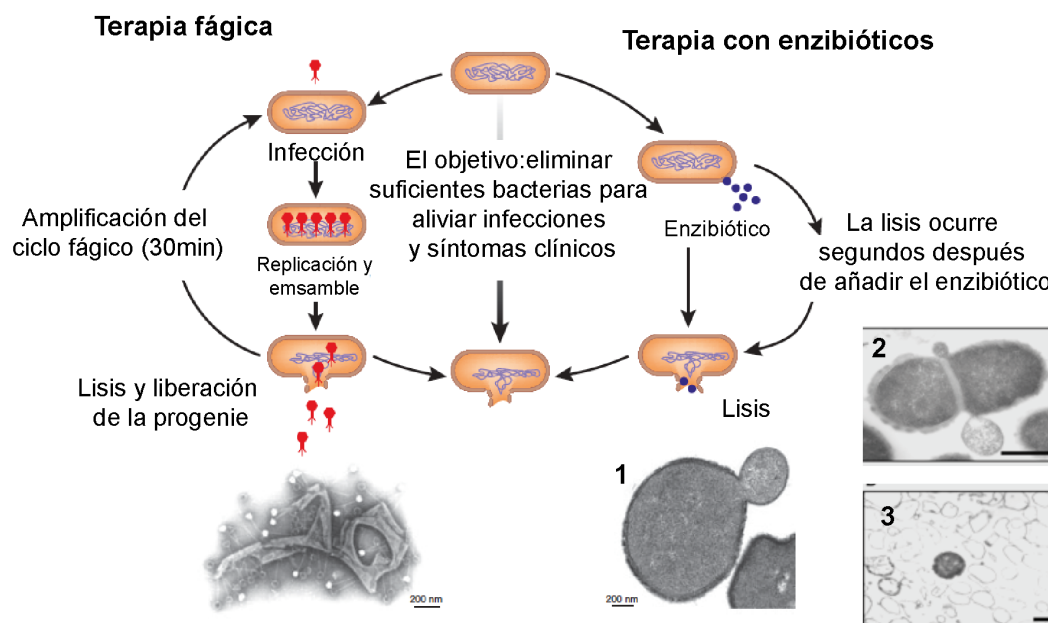


Figura 1.15. Representación esquemática de los pasos en terapia fágica y con enzibióticos. Terapia fágica (a la izquierda) se rige por el ciclo lítico natural de los fagos, el cual ocurre en 30 minutos aproximadamente y lleva tres pasos hasta la liberación de los nuevos fagos (rojo), los cuales infectan nuevas bacterias y amplifican el proceso. La micrografía electrónica muestra partículas víricas adheridas a bacterias lisadas de estreptococo. En comparación la terapia con enzibióticos (a la derecha), es completada en tan sólo dos pasos, en donde la lisina purificada (azul), al contacto inmediato con la superficie de la bacteria, degrada la pared y la bacteria sufre una lisis osmótica. La micrografía electrónica 1 muestra un corte de *Bacillus anthracis* tratado con la endolisina PlyG mostrando una externalización de la membrana citoplasmática (escala de la barra es 200 nm), (modificada de Fischetti *et al.*, 2006. La micrografías 2 y 3 muestra la lisis de *S. pneumoniae* tratado con la endolisina Pal; 2, a un segundo y 3, a 5 segundos del contacto con la endolisina donde sólo quedan fantasmas celulares sin citoplasma Loeffler *et al.*, 2001).

1.2 ESTRATEGIAS DE RECONOCIMIENTO DEL PATÓGENO POR EL HOSPEDADOR

1.2.1 Sistema inmune

Evolutivamente el sistema inmune surgió como respuesta a las infecciones causadas por virus, bacterias, protozoos, hongos y helmintos. Estos patógenos pueden ser responsables de infecciones intracelulares o extracelulares, para las cuales la respuesta inmune debe de ser diferente. De hecho, el sistema inmune ha desarrollado una variedad de respuestas apropiadas para combatir cada tipo de patógeno al mismo tiempo que posee la tolerancia a los componentes del propio organismo.

La eliminación de un patógeno que haya establecido una infección (atravesando las barreras epiteliales del organismo), requiere en primer lugar el reconocimiento por el sistema inmune, y a continuación se debe desarrollar una respuesta adecuada para eliminarlo. El sistema inmune presenta dos tipos de mecanismos, innatos y adaptativos, cuya diferencia principal reside en las estructuras de reconocimiento de los patógenos, ya que los mecanismos efectores de destrucción son esencialmente similares.

Los mecanismos innatos, más primitivos evolutivamente hablando, de acción inmediata, son procesos generalmente inespecíficos de reconocimiento del patógeno, encargados de combatir la infección desde el mismo momento de su inicio y durante sus primeras fases (aproximadamente de 0-5 días) con gran eficiencia. Si estos mecanismos no logran eliminar la infección al menos la mantienen bajo control mientras se desarrollan otro tipo de mecanismos, los adaptativos, que requieren más tiempo para producir la respuesta efectora. La inmunidad innata se basa en la activación del sistema del complemento, activación de la fagocitosis por parte de los macrófagos y la respuesta inflamatoria (mastocitos), que tienen lugar gracias a los receptores innatos para múltiples patógenos. Existen otros mecanismos innatos que no actúan inmediatamente sino que son inducidos, como son la respuesta de la fase aguda, los interferones (especializados en la defensa frente a los virus), las células citolíticas naturales (NK, *natural killer*), las citoquinas y otros mediadores que producen inflamación.

Los mecanismos adaptativos tardan una semana en desarrollarse y poseen mecanismos de reconocimiento al patógeno extremadamente específicos (receptores para antígeno) y presentan memoria. Los linfocitos T son los responsables de la defensa frente a patógenos intracelulares y los linfocitos B en el caso de los patógenos extracelulares, como *S. pneumoniae*. Para combatir los patógenos extracelulares o sus productos, los linfocitos B secretan anticuerpos que se unen específicamente a ellos y actúan como adaptadores entre el patógeno y el mecanismo destructor adaptativo, como es el sistema del complemento, activado por la vía clásica y la fagocitosis dependiente de anticuerpos.

El primer y más elemental mecanismo de defensa frente a las infecciones es el de los epitelios, que constituyen barreras mecánicas, químicas (defensinas y lisozimas) y microbiológicas contra las infecciones. Las mucosas además están bañadas por fagocitos e inmunoglobulinas de tipo A (IgA). Si los patógenos logran atravesar esta barrera y establecer una infección, existen dos mecanismos innatos, preexistentes en el individuo, que actúan inmediatamente para erradicarla: humoral, activando el sistema del complemento, y celular, formado por los fagocitos (monocitos y macrófagos residentes y neutrófilos importados) e inflamocitos.

1.2.2 Sistema del complemento

El sistema del complemento en mamíferos es el mecanismo efector humoral más importante de la respuesta inmune que, junto a los fagocitos, es responsable de la inmunidad innata. El sistema está formado por varias proteínas plasmáticas (es decir, que circulan en sangre que empapa los tejidos) y proteínas de membrana. La activación de estas proteínas es una reacción en cadena que da lugar a la eliminación del patógeno. La eliminación del patógeno puede ser de forma directa, a través de la formación de poros en la membrana del patógeno provocando la lisis o, de forma indirecta, promoviendo la degradación del patógeno por los macrófagos y la inflamación, atrayendo a otras células y moléculas efectoras a la zona infectada.

En la figura 1.16 se esquematizan las vías de activación del complemento reportadas hasta ahora:

1. La vía alternativa, que apunta directamente a la hidrólisis de C3 por el factor B.
2. La vía de las lectinas, iniciada por la unión de la MBL (de sus siglas del inglés: *mannose binding lectin*) a carbohidratos específicos de superficie de la bacteria y asociada a la proteasa MASP (serín-proteasa asociada a MBL).
3. La vía clásica del complemento, activada por la formación de complejos con inmunoglobulinas y por mecanismos independientes de anticuerpos a través del receptor SIGN-R1.

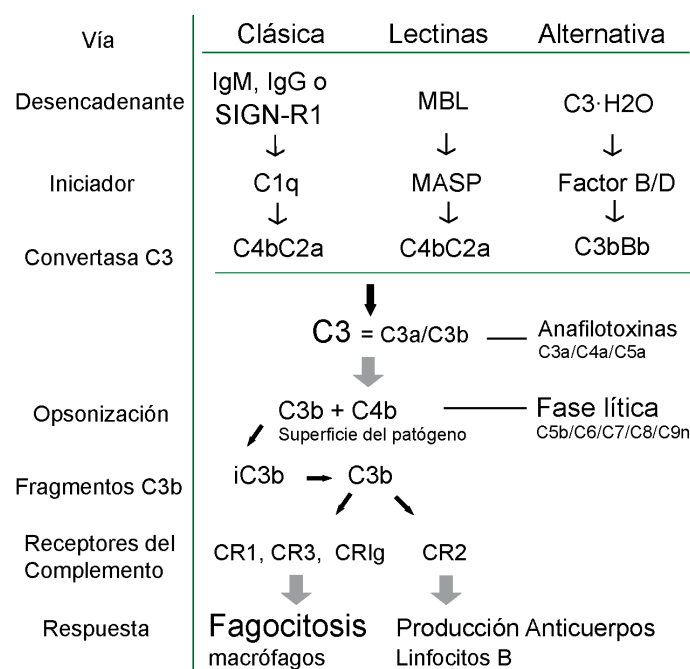


Figura 1.16. Sistema del complemento. Representación esquemática de las diferentes vías de activación y mecanismos efectores resultantes.

La activación del complemento por parte de cada una de las vías genera una convertasa (proteasa) de C3 específica; así C3 es escindida en dos fragmentos solubles, la anafilotoxina C3a y la opsonina C3b. El resultado de esta degradación provoca cambios en la conformación de C3b en donde expone grupos importantes para fijarse a la superficie de la bacteria. Este mecanismo está altamente conservado desde invertebrados a humanos, lo que indica que la formación de C3b es un punto central de alta importancia para la inmunidad innata y adaptativa. A su vez, la formación de C3b induce la formación de la convertasa de C5, que es la responsable del ensamblado de varias proteínas que conllevan a la formación de poros sobre la membrana del patógeno provocando su lisis (fase lítica). De esta forma, las opsoninas provocan lisis directa del patógeno y por otro lado inducen la fagocitosis a través de receptores para productos del complemento y las anafilotoxinas promueven la respuesta inflamatoria (Fearon & Wong, 1983; Walport, 2001a; Walport, 2001b).

La vía alternativa del complemento se inicia cuando el factor B hidroliza a C3 ($C3b \cdot H_2O$), el cual es continuamente producido en el suero en muy bajas cantidades. Como C3b es capaz de unirse covalentemente a la superficie de la membrana del patógeno, éste es protegido de la acción inactivadora del factor I, que hidrata a C3b para generar iC3b (C3b inactivo), C3f, C3c y C3dg. El complejo entre C3b y el factor B (C3bB) es hidrolizado a su vez por el factor D para dar lugar a la convertasa de C3, C3bBb. Este es estabilizado por el factor P (properdina), formándose así más moléculas de C3b, las cuales se unen al complejo C3bBb para formar la convertasa de C5 (C3bBb3bP). Una vez se forma C5b se inicia la vía lítica formadora de poros en las membranas del patógeno (figura 1.16).

La vía de la lectinas es activada por la MBL, la cual reconoce residuos de manosa y N-acetil glucosamina presentes en muchas superficies de patógeno tales como bacterias y hongos (Weis *et al.*, 1998). Cuando la MBL se adhiere a la superficie del patógeno invasor, se activa una proteasa asociada a la MBL, la MASP, la cual es capaz de generar C2a y C4b. Ambas proteínas componen la convertasa de C3 para formar C3a y C3b. De esta forma se induce la actividad lítica del complejo de ataque a la membrana del patógeno gracias a que se forma la convertasa de C5 (C4bC2aC3b) (Ikeda *et al.*, 1987; Tenner *et al.*, 1995). En el caso específico de *S. pneumoniae*, se ha demostrado que la MBP no es capaz de reconocerlo (Neth *et al.*, 2000; Krarup *et al.*, 2005), y que además los niveles de MBP no son inducidos por la infección de dicho patógeno (Kronborg & Garred, 2002; Atkinson *et al.*, 2004).

1.2.3 La vía clásica del complemento

La vía clásica del complemento es el mecanismo más importante requerido para la inmunidad innata y adaptativa contra *S. pneumoniae* (Brown *et al.*, 2002). La vía clásica del complemento es tradicionalmente activada por la unión de la proteína del complemento C1q a la porción Fc de las inmunoglobulinas pentaméricas M (IgM) para la respuesta adaptativa o las inmunoglobulinas G (IgG) en la respuesta innata (figura 1.17). Estos anticuerpos reconocen motivos específicos de superficie del patógeno. La unión de la C1q a los anticuerpos activa a la C1r, la cual inicia una serie de eventos proteolíticos para generar C4b y C2a que al ensamblarse forman la convertasa de C3 de la vía clásica (C4bC2a). La digestión de C3 en un principio da lugar a los fragmentos C3a (anafilotoxina) y C3b, y más tardíamente a C3bi y C3d. C3b, C3bi y C3d que funcionan como ligandos u “opsoninas”, los cuales tienen receptores sobre las células B, fagocitos y células dendríticas foliculares (FDCs). Una vez que las opsoninas son puestas en juego, se activan las diferentes vías efectoras como son la vía o fase lítica descrita anteriormente, la activación de la fagocitosis (Fearon & Wong, 1983), la estimulación de las células B para la producción de anticuerpos específicos (Carroll, 1998) y la retención de antígenos por FDCs (Pozdnyakova *et al.*, 2003), así como también la respuesta inflamatoria en donde los péptidos de C3a causan una contracción involuntaria del músculo, y las células liberan histaminas y se produce inflamación local.

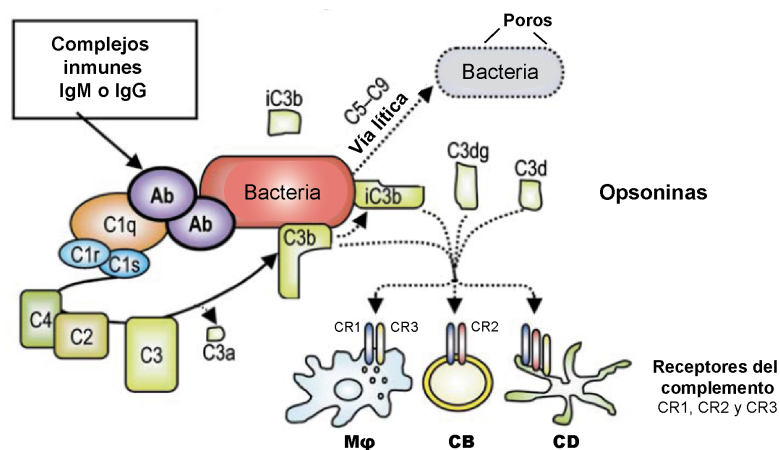


Figura 1.17. Vía clásica del complemento. Los complejos inmunes inducen la formación de las opsoninas C3b, C3bi, C3d, que activan por un lado la vía o fase lítica y por otro lado la fagocitosis, estimulación de la producción de anticuerpos a través de los receptores específicos CR1, CR2 y CR3. Mφ, Macrófagos (fagocitosis); CB, Células B (productoras de anticuerpos); CD, células dendríticas foliculares (retención de antígenos).

1.2.4 Vía clásica del complemento iniciada por SIGN-R1

Los microorganismos encapsulados como *S. pneumoniae* son resistentes a la vía del complemento ya que la cápsula polisacárida limita el acceso a la membrana bacteriana, reduciendo la efectividad de la opsonización (Abeyta *et al.*, 2003). Por ello el hospedador ha desarrollado estrategias alternativas, mediante receptores específicos que reconocen motivos o patrones moleculares dispuestos en la superficie del patógeno, PRRs (de sus siglas en inglés: *pattern recognition receptors*) (Turner, 1992; Avrameas *et al.*, 1996).

Entre estos PRRs figuran lectinas de tipo C, que son capaces de discriminar patrones moleculares de los patógenos y que permiten un reconocimiento diferenciado del patógeno y sus productos (Gordon, 2002). Las lectinas de tipo C involucradas en estos procesos son expresadas por varios tipos celulares como son las células dendríticas (DCs), macrófagos y otras células presentadoras de antígenos (Meyer-Wentrup *et al.*, 2005). En vista de las crecientes evidencias queda demostrado el importantísimo y crítico papel que juegan las lectinas de tipo C en la inmunidad innata, trabajando como receptores sobre las células dendríticas y macrófagos.

SIGN-R1 es una lectina de tipo C y es uno de los ocho genes homólogos de DC-SIGN en humanos (Park *et al.*, 2001; Powlesland *et al.*, 2006). SIGN-R1 es expresada en altas concentraciones por los macrófagos ubicados en la zona marginal del bazo, órgano especializado en capturar patógenos (Humphrey & Grennan, 1981; Kang *et al.*, 2003), tales como *S. pneumoniae* (Lanoue *et al.*, 2004), y los ganglios linfáticos medulares en ratón (Geijtenbeek *et al.*, 2002; Kang *et al.*, 2003). SIGN-R1 es el principal receptor de polisacáridos tales como el dextrano y la cápsula polisacárida de *S. pneumoniae* de varios serotipos (Kang *et al.*, 2004). En ensayos *in vivo* en los que se inactivó la expresión de SIGN-R1, se observó una inhibición total del reconocimiento de cápsula polisacárida, bloqueando la fagocitosis de *S. pneumoniae* por los macrófagos, y dificultando la presentación de las células foliculares, evitando la producción de anticuerpos (Geijtenbeek *et al.*, 2000; Kang *et al.*, 2003;

Kang *et al.*, 2004; Lanoue *et al.*, 2004). Debido a ello, los animales se hicieron más susceptibles a la infección por neumococo.

La falta de efectividad en la inmunidad observada tras desactivar la expresión de SIGN-R1, sugiere que hay una conexión entre SIGN-R1 y la formación de la convertasa de C3. Kang y colaboradores (2006), demostraron que SIGN-R1 es un nuevo activador de la vía clásica del complemento, actuando como receptor sobre los macrófagos reconociendo directamente motivos de la cápsula polisacárida de *S. pneumoniae* a la vez que une C1q. Esta unión provocaría la producción de C4b y C2a, que forman la convertasa de C3, de una manera independiente de anticuerpos y promovería la fijación de C3 a la cápsula de neumococo. Por todo ello, es indudable que SIGN-R1 juega un papel crucial en la resistencia contra organismos encapsulados (Kang *et al.*, 2006) (figura 1.18).

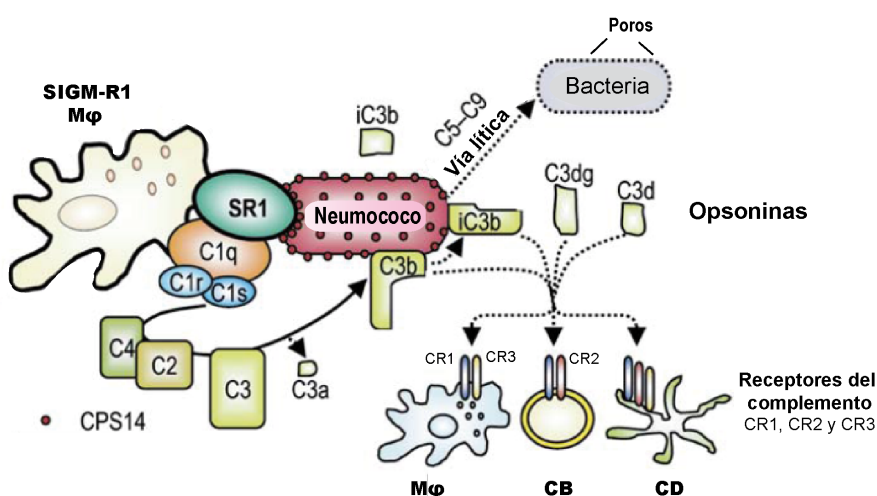


Figura 1.18. Modelo esquemático de la vía clásica del complemento activada por SIGN-R1. En la superficie del macrófago, SIGN-R1 es capaz de unir los polisacáridos de la cápsula de neumococo del serotipo 14 (CPS14), (reemplazando los complejos inmunes usualmente requeridos para la activación de la vía clásica del complemento) para unir a la C1q y ensamblar la convertasa de C3 (C4bC2a). La proteólisis de C3 genera las opsoninas C3b, iC3b y C3dg, que son fijadas a la superficie de la bacteria o la CPS. Las opsoninas generadas de C3, a su vez, permiten la unión a los receptores del complemento sobre los macrófagos (Mφ), células B (CBs) y células dendríticas foliculares (FDCs). Modificada de Kang *et al.*, 2006.

1.2.5 Familia de las lectinas tipo-C

SIGN-R1 pertenece a la superfamilia de las lectinas dependientes de calcio, CTLDs (de sus siglas en inglés: *c-type lectin-like domains*). Todos los miembros de esta familia son proteínas extracelulares, involucrado en procesos de superficie celular, adhesión, reciclado de glicoproteínas del suero y la ya mencionada participación en la respuesta inmune innata (Drickamer & Taylor, 1993). Las CTLDs se caracterizan por tener una región especializada para el reconocimiento de carbohidratos (CRD, de sus siglas en inglés: *carbohydrate-recognition domain*), y en la actuación depende de calcio. Por comparación de secuencias primaria, las lectinas de tipo C, están altamente conservadas en cuanto a los residuos involucrados en la estabilización del átomo de calcio. Los CRD de esta familia de lectinas muestran estructuras globulares cuyo plegamiento fue descrito por primera vez en 1991 (Weis *et al.*, 1991). La especificidad de cada miembro de esta familia viene dada por cambios sutiles en pequeñas regiones que aportan flexibilidad (Drickamer, 1999).

Powlesland y colaboradores (2006), examinaron la secuencia genómica de ratón en aras de identificar los genes homólogos de humanos para la DC-SIGN (de sus siglas en inglés: *Dendritic Cell Specific Intercellular adhesions molecules-3 (ICAM-3)-Grabbing Nonintegrin*), encontrando 8 genes relacionados, con especificidades diferenciadas. DC-SIGN ha sido identificada como un potente receptor de patógenos y como una molécula de adhesión que facilita la unión entre las células T a la células dendríticas y ubicadas sobre las células endoteliales del hígado, nódulos linfoides y placenta. Este grupo de receptores ha sido considerado de gran interés debido a su habilidad para unir virus y parásitos tales como el VIH y subsecuente presentación de antígenos a las células T, facilitando la eficacia de la infección (Pohlmann *et al.*, 2001). Tanto DC-SIGN como su receptor relacionado, DC-SIGNR, son capaces de unir glicoproteínas de virus con alto contenido de manosa (de 8 a 9 residuos).

Los genes homólogos humanos identificados en ratón reflejan una gran versatilidad, siendo capaces de unir cadenas glicánicas de diferente naturaleza como las terminadas en ácido siálico o GluNAc, entre otros. En la figura 1.19 se esquematiza la distribución de dichas proteínas, en donde se destaca el CRD presente en todas ellas en el extremo C-terminal. A este dominio le siguen una región con repeticiones de 23 aminoácidos con asociación tetramérica sobre la membrana (Feinberg *et al.*, 2005), una región transmembrana y finalmente un péptido señal en el extremo N-terminal citoplasmático, cuya función está involucrada en los procesos de internalización del receptor (Guo *et al.*, 2004). Las SIGN-Rs se expresan en diferentes tejidos, por lo que se postula que la especificidad es dependiente del tejido (Park *et al.*, 2001).

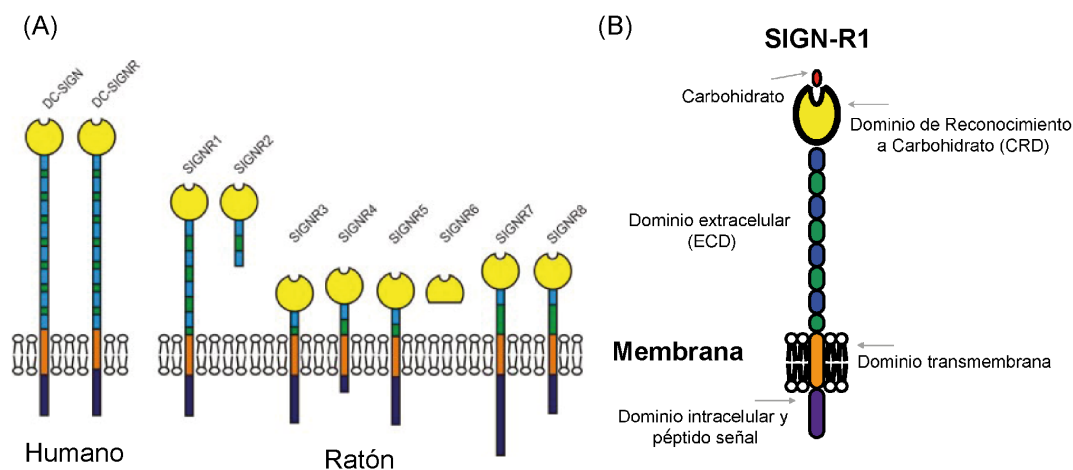


Figura 1.19. Representación esquemática de la distribución del receptor DC-SIGN y sus homólogos.

1.2.6 SIGN-R1 y su participación en la inhibición del proceso inflamatorio

Los anticuerpos pueden considerarse como moléculas adaptadoras que proporcionan un vínculo entre los mecanismos de defensa, celulares y humores. La función de los anticuerpos es la formación de complejos inmunes que resultan de la unión a los antígenos. Estos complejos inmunes pueden activar múltiples mecanismos efectores dando lugar a la retirada y destrucción de un agente extraño al organismo o patógeno. La inmunoglobulina de tipo G (IgG), está compuesta por tres proteínas independientes conectadas a través de un conector flexible, que a su vez permite diferenciar a los IgGs en dos dominios o regiones, Fab y Fc. La región Fab expresa sitios de unión

específicos para el reconocimiento de los antígenos. La región Fc expresa sitios que son reconocidos por diversos receptores responsables de activar los mecanismos efectores. El conector entre ambos dominios, la región bisagra, le confiere al IgG una gran flexibilidad, permitiendo así que el Fab pueda acceder al antígeno diana, y al mismo tiempo permite que el Fc pueda ser reconocido por los receptores de Fc (figura 1.20-A).

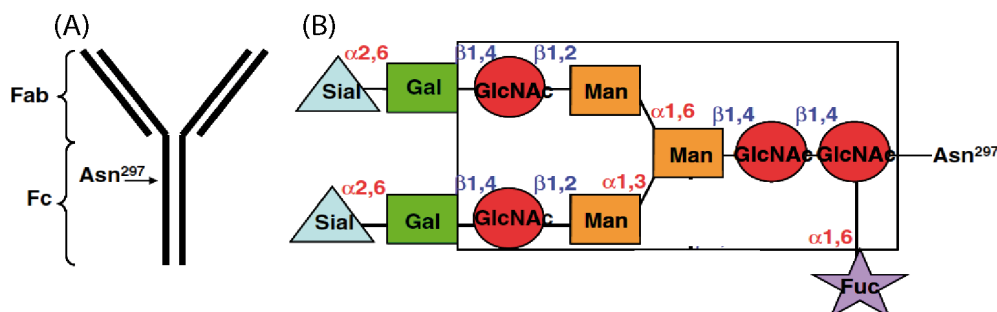


Figura 1.20. Representación esquemática de la inmunoglobulina G (IgG). (A) El IgG está compuesto por dos dominios, la región Fab y la región Fc. Se señala la glicosilación en Asn²⁹⁷ con una flecha. (B) Glicosilación característica de doble cadena, constituida por un centro de 7 azúcares (enmarcado), seguidos por Gal y Sial en posición terminal. El texto en rojo y azul muestra los enlaces entre los azúcares. Fuc: fucosa, GlcNAc: N-acetilglucosamina, Man: manosa, Gal: galactosa, Sial: ácido siálico.

Los receptores de la región Fc (FcγRs) son proteínas de membrana expresadas constitutivamente en diversas células de sistema inmune como células B, células T, macrófagos, neutrófilos, células cebadas, etc. El reconocimiento de los complejos inmunes por parte de los FcγR, pueden iniciar tanto la activación como la inhibición de diversos mecanismos efectores (Chan *et al.*, 1994). Las respuestas que conlleva el reconocimiento de los Fc por sus receptores celulares incluyen la activación o inhibición de la citotoxicidad dependiente de anticuerpo (Fanger *et al.*, 1989), la presentación de antígenos y la regulación de la producción de anticuerpos (Fridman, 1993; Guyre *et al.*, 2001), la fagocitosis (Anderson *et al.*, 1990) y la secreción de mediadores involucrados en procesos inflamatorios (Anegón *et al.*, 1988). En humanos hay tres clases de FcγRs dependiendo de la afinidad por el Fc, receptores de alta afinidad (K_D 10^{-8} M) FcγRI y receptores de baja afinidad FcγRII y FcγRIII (K_D 10^{-6} – 10^{-7} M) (Ravetch & Kinet, 1991). La mayoría de los Fc de los IgGs están glicosilados y su reconocimiento está modulado por la naturaleza de la cadena glicánica. Así por ejemplo, dependiendo de la longitud de dicha cadena, se promueven ciertos cambios de conformación que favorecen o inhiben su interacción con los diversos receptores (Krapp *et al.*, 2003).

Las IgG son las inmunoglobulinas más abundantes en el suero humano, teniendo una gran relevancia en el tratamiento de enfermedades autoinmunes como, la enfermedad de Kawasaki, el síndrome de Guillain-Barré, y neuropatías crónicas desmielinizantes, gracias a las propiedades antiinflamatorias que poseen (Nimmerjahn & Ravetch, 2007). La propiedad antiinflamatoria de las IgG, está asociada a la región Fc de estos anticuerpos. Sólo los Fc que están glicosilados y cuya cadena glicánica termina en ácido siálico (α-2,6) (figura 1.20-B), son capaces de promover la respuesta antiinflamatoria (Anthony *et al.*, 2008a).

Recientemente se ha reportado que estos Fc-sialilados requieren de SIGN-R1 como receptor principal y crucial para iniciar una cascada de eventos que producen la inhibición de la inflamación

(Anthony *et al.*, 2008b). Cuando los Fc-sialilados son reconocidos por SIGN-R1, los macrófagos reguladores segregan moléculas solubles que afectan a los macrófagos efectores, previamente activados por los complejos inmunes antígeno/anticuerpo (ubicados en los tejidos infectados e inflamados). Los macrófagos efectores son entonces inducidos a expresar receptores que inhiben el proceso inflamatorio (Anthony & Ravetch, 2010), (figura 1.21). Por esta razón la inyección intravenosa de los Fc-sialilados es un potente tratamiento para pacientes con enfermedades auto inmunes en donde los procesos inflamatorios son un grave problema.

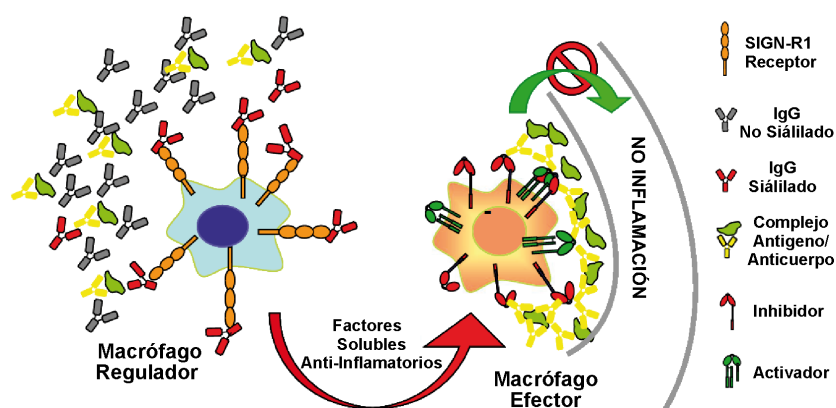


Figura 1.21. Modelo de la actividad antiinflamatoria del Fc sialilado α -2,6. Los anticuerpos sialilados entran en contacto con SIGN-R1 (lectina receptora) en la superficie de los macrófagos reguladores (azul) en la zona marginal del bazo. Esta interacción induce un programa celular que produce la secreción de moléculas solubles que interfieren con el mecanismo de los macrófagos efectores (naranja) que se encuentran en los tejidos en donde se está produciendo la inflamación y los complejos inmunes han sido reconocidos. Cuando los macrófagos efectores entran en contacto con los mediadores anti-inflamatorios se dispara la expresión de receptores inhibitorios en su superficie y de esta forma el proceso inflamatorio se bloquea.

OBJETIVOS

2 OBJETIVOS

En este trabajo se describe el estudio estructural del reconocimiento de los motivos moleculares que recubren la superficie celular de *S. pneumoniae*, específicamente la pared bacteriana y la cápsula polisacárida, en aras de ampliar los conocimientos implicados en la relación patógeno/hospedador y sus posibles aplicaciones terapéuticas.

Las enzimas líticas producidas por los bacteriófagos (endolisinas) han sido parcialmente estudiadas a nivel genético y bioquímico. Sin embargo, no existe un conocimiento estructural detallado de estas enzimas. El uso de las endolisinas como enzibióticos supone una alternativa prometedora en la prevención y el tratamiento de las infecciones producidas por neumococo, gracias a su especificidad y a la baja probabilidad de que den lugar al desarrollo de resistencias. Su eficacia como agentes antibacterianos se encuentra bien documentada en la literatura reciente y son cada vez más los grupos de investigación interesados en estudiar las propiedades, funcionalidad y posibles aplicaciones de estas endolisinas. Bajo este contexto, es interesante el estudio estructural y funcional de la endolisina Cpl-7 que, además de ser una proteína lítica de alta efectividad, tiene un módulo de unión a la pared celular desconocido y que está presente en bacteriófagos de otros patógenos distintos de neumococo.

De forma paralela, se ha abordado el estudio estructural de un receptor del sistema inmune del hospedador recientemente descrito, SIGN-R1, el cual es capaz de activar mecanismos efectores como la vía clásica del complemento, cuando el neumococo se halla circulando en sangre, siendo capaz de reconocer el mayor factor de virulencia de neumococo, la cápsula polisacárida. El estudio detallado de SIGN-R1 a nivel estructural, proporcionaría grandes avances relativos al conocimiento de la especificidad para el entendimiento de la relación patógeno/hospedador. Además, colateralmente, SIGN-R1 juega un papel esencial en la respuesta antiinflamatoria a través de la unión de otras moléculas efectoras, como los Fc-sialilados.

Así, los objetivos concretos que se proponen en este trabajo pueden sintetizarse en:

1. Caracterización estructural de los mecanismos de hidrólisis y reconocimiento de la pared de neumococo por la endolisina Cpl-7.
2. Caracterización de las bases estructurales del reconocimiento de neumococo por el receptor SIGN-R1.
3. Caracterización estructural de los mecanismos de reconocimiento y activación de las proteínas del sistema inmune por SIGN-R1.

TÉCNICAS EXPERIMENTALES

3 TÉCNICAS EXPERIMENTALES

3.1 VISIÓN GENERAL DEL PROCESO EN CRISTALOGRAFÍA DE RAYOS X

La cristalografía de rayos X es la disciplina científica más potente para la determinación de la estructura tridimensional de la materia ordenada, tanto de naturaleza inorgánica como orgánica. Nos permite obtener una imagen tridimensional a escala atómica del material cristalizado. Aplicado al campo de las proteínas, proporciona toda la información estructural, como el plegamiento y detalles de un gran significado físico-químico (distancias, ángulos de enlace, detalles de los sitios activos), lo que constituye información clave para entender la función biológica y los procesos en los que están involucradas. Además, con la cristalografía posible estudiar desde sistemas sencillos hasta los más complejos, como la estructura de un virus o los grandes agregados macromoleculares, como es el caso de los ribosomas. La resolución estructural de las proteínas mediante cristalografía de rayos X consta de las siguientes etapas:

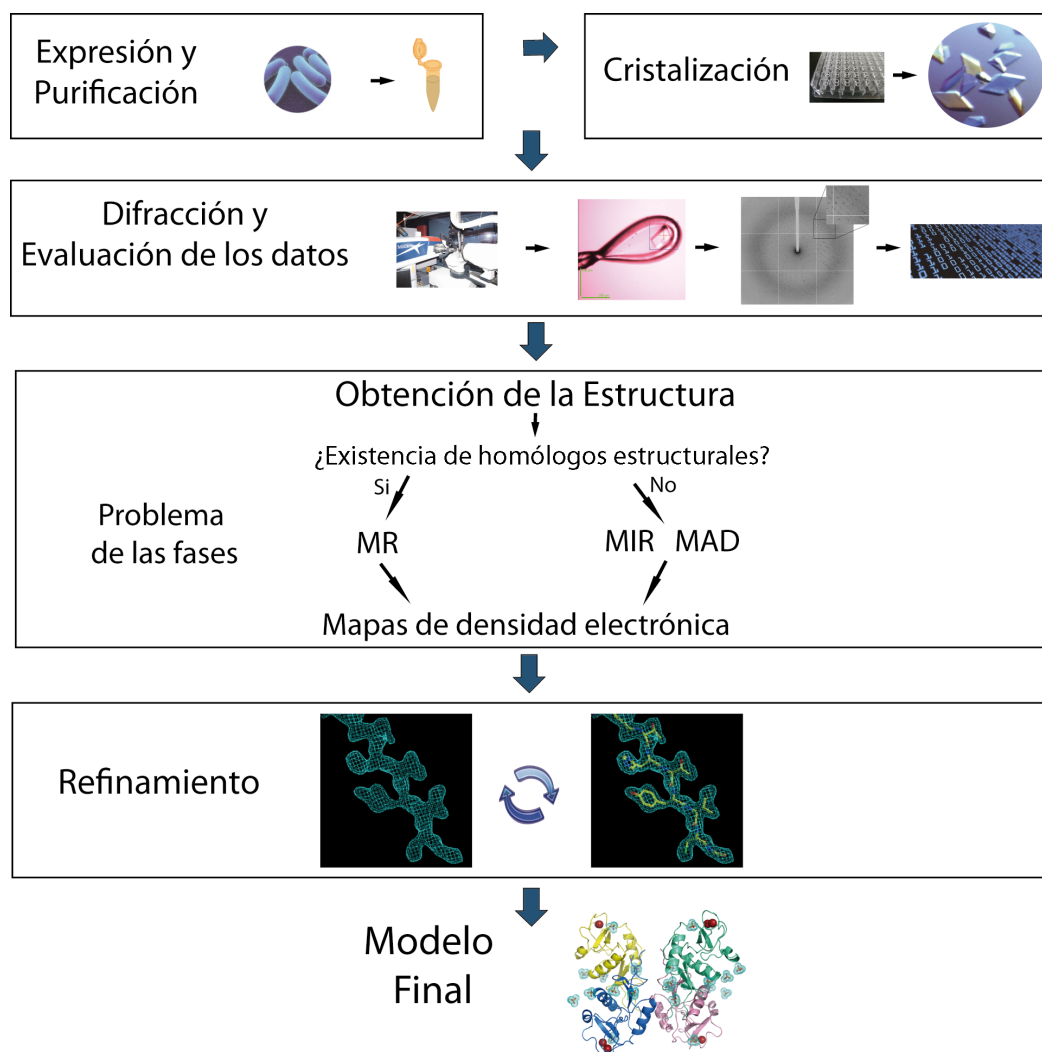


Figura 3.1. Esquema general de las etapas en el proceso de resolución estructural de proteínas mediante cristalografía de rayos X.

3.2 CONSIDERACIONES TEÓRICAS

3.2.1 Cristalización

La determinación de la estructura tridimensional de cualquier molécula por difracción de rayos X depende, en primer lugar, del éxito en la obtención de monocristales con tamaño adecuado para someterlos a difracción (McPherson, 1989; McPherson & DeLucas, 1999).

Un cristal se puede considerar como una red de unidades elementales idénticas, denominadas cada una de ellas celdilla unidad. La repetición por translación en las tres direcciones del espacio de este elemento estructural genera el volumen del cristal. El cristal está formado por celdillas unidad. La celdilla unidad a su vez, presenta en su interior una unidad mínima denominada unidad asimétrica, cuya repetición mediante ciertas operaciones de simetría genera el contenido de la celdilla unidad. Dicha celdilla unidad está definida por tres ejes (a , b , c) y tres ángulos interaxiales (α , β , γ) (figura 3.2.) que se usan como sistema de referencia para las coordenadas atómicas. La combinación de los elementos de simetría contenidos en la celdilla elemental designa el grupo espacial, que describe el patrón de repetición que presentan una o varias moléculas (unidad asimétrica) en el espacio tridimensional.

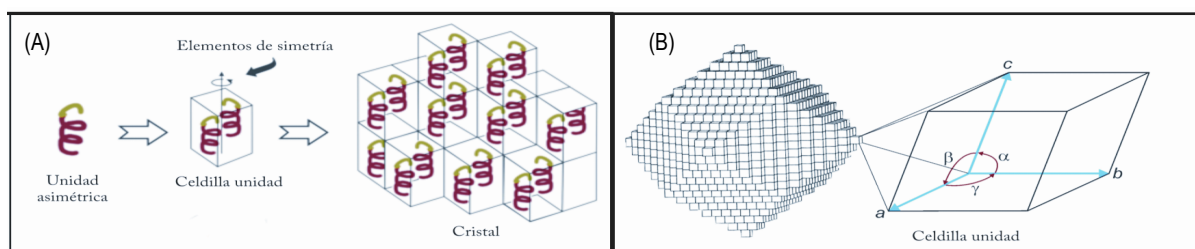


Figura 3.2. (A) La unidad asimétrica se repite a sí misma mediante los elementos de simetría, reproduciendo la celdilla unidad. (B) El apilamiento de celdillas unidad forma el cristal.

Los cristales de proteína difieren bastante de los cristales convencionales de moléculas pequeñas (como los compuestos orgánicos en general) o sales inorgánicas (como la sal común), ya que las características de los primeros están condicionadas por la propia naturaleza de estas macromoléculas. Así:

Al estar formados por moléculas con quiralidad, se describen exclusivamente con 65 grupos espaciales (de los 230 posibles), es decir, aquellos en los que no existen las operaciones de repetición que impliquen centros o planos de simetría.

Debido al tamaño de las proteínas y su esfera de solvatación, el volumen de la celdilla unidad suele ser bastante grande, en comparación con los cristales de moléculas pequeñas.

Como las macromoléculas empaquetan con baja densidad, al ordenarse en el cristal dejan grandes canales constituidos por moléculas de agua desordenadas y el contenido en solvente suele ser elevado (50-80%) (Matthews, 1968).

Estos factores determinan que la fuerza de interacción entre las moléculas sea débil, de manera que los cristales son muy frágiles, sensibles a la radiación y difíciles de manipular. Por otra

parte, el elevado contenido en solvente desordenado hace que los datos de difracción que se obtienen sean generalmente muy débiles. Este efecto puede ser especialmente acusado cuando en la molécula existen segmentos de alta movilidad, lo que provoca cierto desorden interno en el cristal que luego se ve reflejado con la pérdida de cierta visibilidad estructural. Sin embargo, el alto contenido en solvente de los cristales de proteína también tiene aspectos positivos, ya que la presencia de una esfera de solvatación permite a la proteína estar correctamente plegada, y la presencia de grandes canales posibilita la inclusión en el cristal de átomos pesados, sales y moléculas, para obtener estructuras de complejos proteína-ligando (sustratos, productos, cofactores o inhibidores) a partir del cristal de proteína nativa.

Por tanto, el primer objetivo en la resolución de la estructura de una proteína, por cristalografía de rayos X, es la obtención de un monocristal que tenga el mayor tamaño (mayor volumen dispersor) y la menor mosaicidad (parámetro que nos informa del orden interno que posee un cristal) posibles. Cuanto mayor desorden interno, mayor será la mosaicidad, lo que se traduce en datos de difracción de baja calidad.

3.2.2 Principios básicos de la cristalización

Para la producción de monocristales de tamaño adecuado es fundamental disponer de una muestra pura y homogénea. Durante el proceso de cristalización, las moléculas en solución, en presencia de agentes precipitantes suaves, son conducidas a un estado de sobresaturación (termodinámicamente metaestable) que posteriormente deriva en la aparición de una fase cristalina en donde las moléculas son llevadas a un equilibrio entre las fases sólida y líquida (Saludjian *et al.*, 1992) (figura 3.3-A).

El proceso de cristalización se puede dividir en dos fases: la fase de nucleación (formación de los primeros agregados ordenados) y la fase de crecimiento del cristal (en donde el conjunto de interacciones atractivas entre las moléculas son máximas, y las repulsivas mínimas).

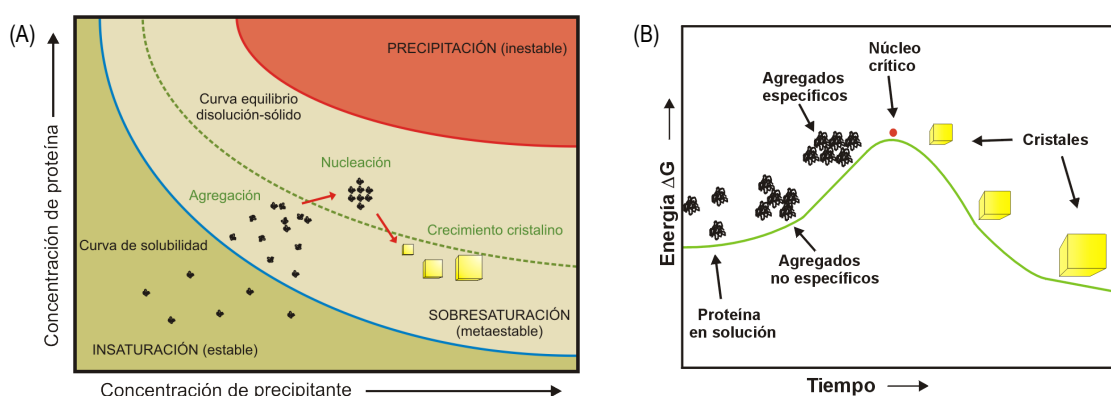


Figura 3.3. (A) Descripción esquemática de un diagrama bidimensional de solubilidad en el que se ilustra el cambio de concentración de proteína con la concentración de agente precipitante. (B) Diagrama de energía de la cristalización. Imagen extraída de la web <http://www-structured.cimr.cam.ac.uk/Course/Crystals>.

Durante este proceso hay que superar una barrera energética, como ocurre en cualquier reacción química (figura 3.3-B). En el caso de las proteínas, la superación excesiva del punto de

saturación, sobre todo si sucede rápidamente, conlleva con gran frecuencia a la precipitación de la proteína en forma de agregados amorfos, no ordenados, que no son adecuados para los estudios de difracción. Por el contrario, si la sobresaturación es leve se alcanza un estado metaestable que no da lugar a núcleos cristalinos (a menos que se le aporte energía al sistema), pero que sería la zona ideal para el crecimiento de cristales sin la aparición de nuevos núcleos. Por tanto, la formación de cristales requiere alcanzar lentamente un grado limitado de sobresaturación, para que se formen pocos núcleos, de manera que la masa proteica se invierta en el crecimiento de los cristales y así éstos adquieran un tamaño adecuado.

3.2.3 Métodos de cristalización

Existen varios métodos de cristalización: difusión de vapor, contradifusión, diálisis o a través de interfase. De todos ellos, el método de difusión de vapor es el utilizado con mayor frecuencia y presenta varias modalidades: gota sentada y gota colgante.

El método de difusión de vapor consiste en equilibrar una gota, en la que hemos mezclado la solución de proteína con la solución precipitante, frente a un reservorio que contiene únicamente solución precipitante. Debido a la diferencia de concentración que existente entre la gota y el reservorio, se establece un proceso de difusión de vapor entre ambos que genera la transferencia de componentes volátiles (esencialmente agua) desde la gota hacia el reservorio. De esta forma, la concentración de proteína y agente precipitante de la gota va aumentando lentamente hasta alcanzar el equilibrio. A lo largo de este proceso de concentración de la gota podemos llegar a alcanzar un estado de sobresaturación de la proteína y obtener, por tanto, cristales (figura 3.4).

Para montar los experimentos de cristalización en gota colgante se mezcla la proteína con el reservorio (gota) sobre un cubreobjetos de vidrio siliconado, se le da la vuelta y a continuación se coloca sobre el pocillo de una placa de cristalización, en cuyo interior hemos añadido la solución precipitante (reservorio). El borde del pocillo se encuentra impregnado con grasa de vacío, de manera que al colocar el cubreobjetos se logra que la gota y reservorio queden aislados del exterior y se establezca un equilibrio de vapor entre ambos.

En el caso de experimentos de cristalización en gota sentada, se mezcla la solución de proteína con el reservorio sobre un pequeño soporte dentro del pocillo, que contiene la solución precipitante, se sella con cinta adhesiva, y se logra que la gota y reservorio queden aislados y se establezca un equilibrio de vapor entre ambos.

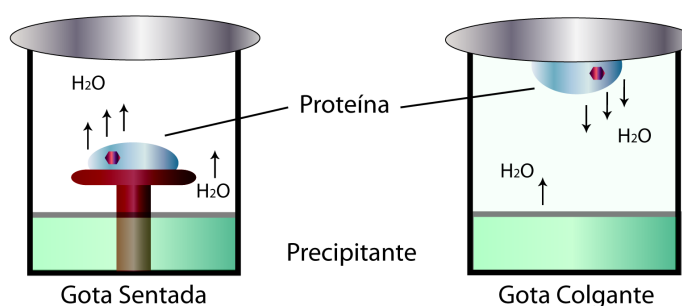


Figura 3.4. Método de difusión de vapor. En el sistema, una vez sellado, se establece un equilibrio de vapor que tiende a igualar las concentraciones entre la gota y el reservorio.

3.2.4 Búsqueda de las condiciones de cristalización

Cuando abordamos la cristalización de una proteína nueva, en general carecemos de información acerca de las condiciones que van llevar a la proteína al estado de sobresaturación antes mencionado. Por tanto, la estrategia a seguir inicialmente consiste en probar sistemáticamente una amplia gama de condiciones en las que se varían todos los parámetros posibles. Para llevar a cabo la búsqueda inicial de la solución precipitante se utilizan baterías de soluciones preparadas comercialmente, entre las que destacan Crystal Screen I, Crystal Screen II, Crystal Screen Lite, SaltRx e Index (Hampton Research, CA, USA); JCSG+Suite y PACT+Suite (QIAGEN) y Precipitant Synergy (Emerald BioSystems).

Hoy en día contamos con técnicas de alto rendimiento que, mediante el uso de robots, permiten ensayar un gran número de condiciones en poco tiempo y además con pequeñas cantidades de proteína. El sistema de los robots Gilson-Innovadyne semiautomatizado nos permite hacer ensayos de cristalización sobre una placa de 96 pocillos por el método de difusión de vapor en gota sentada, utilizando solo 250 nL de solución de proteína por condición y 70 μ L de solución precipitante o reservorio.

Solución precipitante

Agente precipitante. Generalmente se utilizan polímeros, como el polietilenglicol (de diferente masa molecular) y sales, y con menos frecuencia compuestos orgánicos como alcoholes. Además de la naturaleza del precipitante también se varía su concentración.

pH. Generalmente se barre un rango de pH desde a 4.0 a 8.5, así como la naturaleza del tampón.

Aditivos. Se utilizan sales, alcoholes, azúcares, detergentes y otras moléculas de pequeño tamaño y a baja concentración.

Solución de proteína

En cuanto a la muestra de proteína, la concentración de partida que suele utilizarse es del orden de 10 $\text{mg}\cdot\text{mL}^{-1}$, concentración que puede variarse en función de la solubilidad y/o estabilidad de la muestra en solución. Una proteína muy soluble generalmente requerirá trabajar a una

concentración mucho mayor para llevarla al estado de sobresaturación, así como esta solubilidad y/o estabilidad pueden verse fuertemente afectadas por la naturaleza del tampón, el pH y la presencia de otros componentes, como sales. En principio es conveniente utilizar tampones a una concentración baja (10-50mM) y evitar aquellos tampones o compuestos que puedan formar cristales y dificultar el proceso de cristalización de la proteína en estudio.

Optimización de cristales de proteína

Una vez encontradas las condiciones iniciales de cristalización, procedemos a su escalado y optimización, variando ligeramente la concentración de proteína, la concentración del agente precipitante, el pH, la presencia de aditivos y detergentes, la temperatura, la relación de los volúmenes de las distintas soluciones mezcladas en la gota (es decir, la cinética de formación del cristal) y el método de cristalización, con el objetivo de mejorar el tamaño y calidad de los cristales de proteína según las necesidades del sistema.

3.2.5 Cristalización de complejos proteína-ligando

Hay diferentes métodos para obtener cristales de proteína-ligando. A este respecto, Hassell y colaboradores (2007) proponen cuatro estrategias a seguir, consistentes en (i) inmersión, (ii) cocrystalización, (iii) coexpresión de la proteína con el ligando y (iv) uso del ligando de interés durante la purificación. De todas ellas, la inmersión y la cocrystalización son las estrategias iniciales a seguir para obtener un complejo proteína-ligando.

Cocrystalización: Esta técnica consiste en la cristalización de la proteína en presencia del ligando de interés. Generalmente, esta estrategia se prueba en primera instancia en las condiciones de cristalización de la proteína nativa (sin ligando), aunque la presencia del ligando puede afectar a dichas condiciones o incluso impedir la cristalización, en cuyo caso habría que buscar unas condiciones de cristalización nuevas para el complejo con el ligando, siguiendo el mismo protocolo explicado anteriormente.

Inmersión (*soaking*): Esta técnica consiste en la inmersión de los cristales de proteína ya formados en una solución con ligando. Dado que los cristales de proteína suelen presentar empaquetamientos con grandes canales ocupados por solvente, los ligandos, que son moléculas muy pequeñas en comparación con estos canales, pueden difundir con facilidad y unirse a la proteína siempre que el sitio de unión se encuentre accesible. Estos ensayos presentan algunas limitaciones en tanto que el sitio de unión del ligando debe de estar accesible o incluso, en algunas ocasiones, la unión del ligando puede afectar al empaquetamiento cristalino generando una pérdida del orden interno del cristal.

3.3 TÉCNICAS CRISTALOGRÁFICAS Y RESOLUCIÓN ESTRUCTURAL

La biología estructural tiene como fin último comprender los procesos biológicos en términos de la física y la química de las macromoléculas que participan en ellos. Una de las diferencias

esenciales entre la química de los sistemas vivos y los inertes es la gran complejidad estructural de las macromoléculas biológicas. Por tanto, la química de la vida no se podrá desvelar sin antes conocer a resolución atómica, o cercana a la atómica, la estructura de las macromoléculas biológicas, especialmente las proteínas (Brändén & Tooze, 1991), que son las verdaderas máquinas de la vida. Así, comprender la estructura de cada proteína es comprender en que consiste la vida (Max Perutz, Premio Nobel de Química, 1962).

3.3.1 Medidas del espectro de difracción

La difracción de rayos X es el fenómeno físico fundamental a través del cual se manifiesta la interacción de esta radiación con la materia cristalina. Cuando un haz de rayos X incide sobre un cristal, éste interacciona con los electrones de los átomos que lo componen, haciéndolos vibrar acopladamente con las variaciones periódicas de su campo eléctrico. Los electrones oscilantes se convierten, de esta manera, en focos de nueva radiación X que se emite en cualquier dirección, fenómeno que se denomina “dispersión”. Estos rayos X dispersados por los electrones interfieren entre sí, pudiendo dar interferencias destructivas. Sin embargo, como los átomos en un cristal están ordenados de un modo regular y periódico, también ocurre que en determinadas direcciones las ondas interfieren constructivamente, reforzándose, dando lugar al fenómeno de dispersión cooperativa o difracción. Es en este caso de interferencia constructiva en el que se cumple la Ley de Bragg a partir de la cual podemos describir el fenómeno de la difracción:

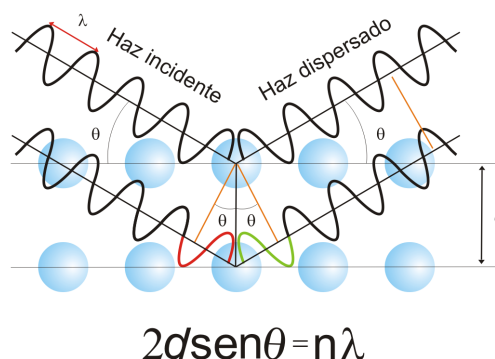


Figura 3.5. Ley de Bragg. La onda que incide en el plano inferior se retrasa un número entero de longitudes de onda y sale en fase con la onda superior, reforzándose la señal.

El hecho de utilizar radiación X se debe a que el fenómeno de difracción requiere que la longitud de onda de la radiación electromagnética sea del mismo orden de las distancias entre los focos dispersores, los átomos del cristal.

El modo en que la cristalografía de rayos X llega al conocimiento de la estructura cristalina de cualquier material puede comprenderse mejor si se compara con el símil que representa la observación de un objeto a través de un microscopio óptico. En éste, la radiación dispersada por el objeto en estudio se recombina de nuevo, a través de un sistema de lentes, para dar lugar a la imagen aumentada del objeto dispersor. Sin embargo, no existe ninguna lente que sea capaz de focalizar los

rayos X dispersados por los átomos. Para ello, las técnicas cristalográficas hacen uso de cálculos matemáticos denominados “síntesis de Fourier”, con los que se consigue “focalizar” los rayos X.

3.3.2 Fuentes de radiación y equipos de difracción

En los experimentos de difracción se hace incidir un haz monocromático de rayos X con una longitud de onda de alrededor de 1 Å sobre un monocristal de la muestra en estudio.

El espectro de difracción de rayos X contiene un patrón de intensidades correspondientes a los haces difractados por el cristal, y la recogida de datos de difracción consiste en la medida de dichas intensidades y su interpretación espacial respecto al sistema de referencia del cristal. El diseño experimental implica la toma de imágenes sucesivas en un detector plano situado en la perpendicular del haz de rayos X, para distintas posiciones angulares del cristal, hasta obtener un conjunto de datos completo. Esto se consigue mediante oscilaciones simples del cristal (geometría Phi) o bien mediante una combinación de giros más complejos (geometría Kappa) (figura 3.6).

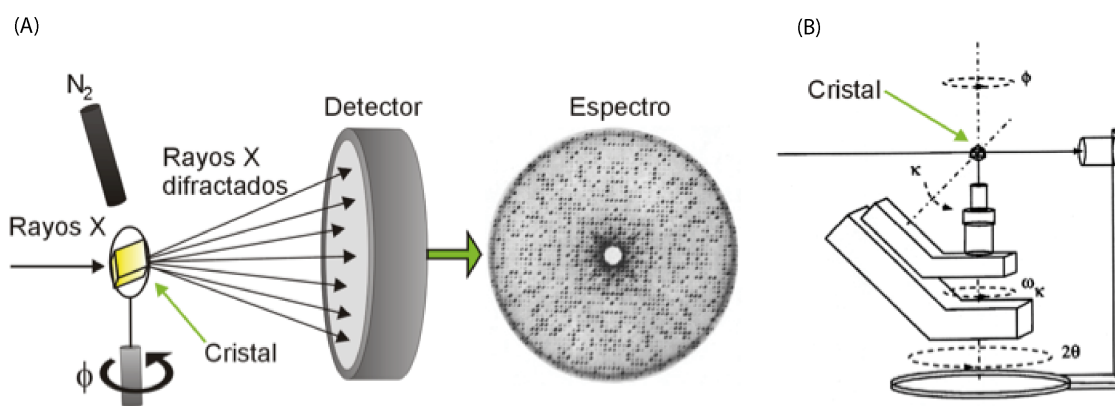


Figura 3.6. (A) Esquema de un experimento de difracción con geometría Phi. (B) Esquema de un difractómetro con geometría Kappa. La estrategia de recogida de los datos de difracción y la elección de la fuente de radiación dependerá del tipo de datos de difracción que sean necesarios para poder abordar la resolución estructural, pudiéndose realizar en un difractómetro con fuente convencional o en una estación de radiación sincrotrón.

Generador de ánodo rotatorio: Los difractómetros utilizados en cristalografía de proteínas usan como fuente de rayos X un generador de ánodo rotatorio (figura 3.7). Este generador permite suministrar un alto voltaje al tubo de alto vacío de rayos X en donde los electrones de un filamento incandescente se desprenden e inciden sobre el ánodo metálico en forma de cilindro. De esta forma, electrones de los átomos del ánodo se excitan pasando a niveles superiores de energía que a continuación abandonan para volver al estado fundamental liberando el exceso de energía en forma de radiación X. Convencionalmente, se utiliza un ánodo de cobre, el cual genera radiación X de 1.5418 Å de longitud de onda, correspondiente a la línea K α del cobre.

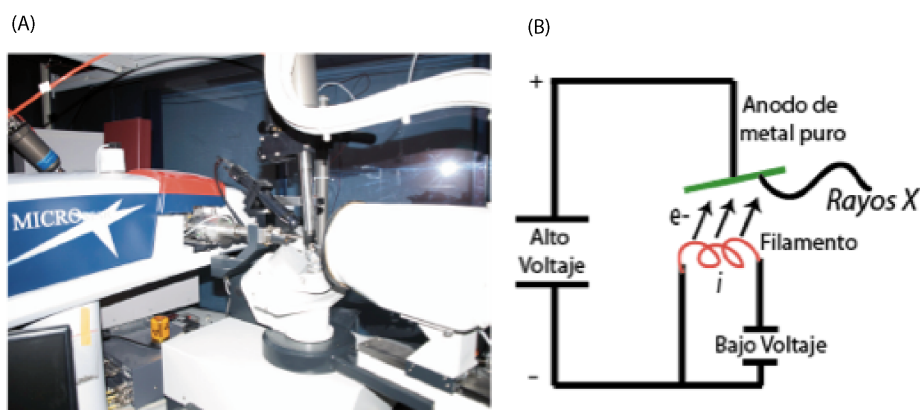


Figura 3.7. Generador de ánodo rotatorio. (A) Equipo Brüker-Nonius perteneciente al Departamento de Cristalografía y Biología Estructural del Instituto de Química-Física Rocasolano (CSIC). (B) Esquema de la generación de rayos X con ánodo de cobre.

Radiación sincrotrón: La radiación sincrotrón es la fuente de radiación X más intensa que existe. El sincrotrón consta de un anillo de cientos de metros de longitud por el que se hacen circular electrones o positrones a velocidades relativistas. Cuando a este haz de partículas se le obliga a cambiar la dirección de su trayectoria mediante imanes, para poder seguir la curvatura del anillo, se produce la emisión de una radiación muy energética y tangencial a la dirección seguida por los electrones denominada radiación sincrotrón. La radiación emitida muestra un espectro continuo de longitud de onda muy amplio, en el que se encuentran incluidos los rayos X. (figura 3.8).

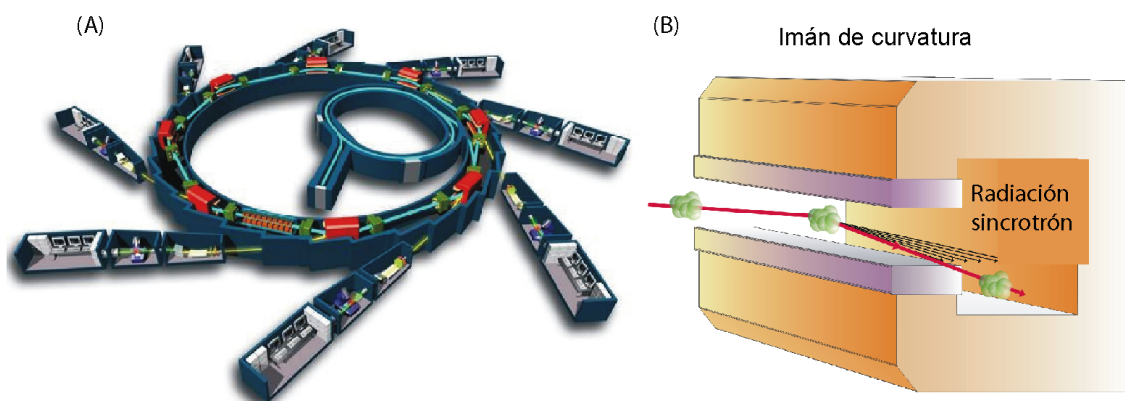


Figura 3.8. Fuente de radiación sincrotrón. (A) Esquema de un sincrotrón. (B) Imán de curvatura utilizado para curvar la trayectoria de las partículas en el interior del anillo.

Si bien las características técnicas de un difractómetro con fuente de radiación convencional le confieren una gran versatilidad y un bajo coste a la hora de diseñar y realizar una estrategia de recogida de datos, la radiación sincrotrón tiene como ventajas fundamentales su mayor brillo (mayor intensidad y más concentrada) y la posibilidad de sintonizar la longitud de onda de los rayos X. Dado que el fenómeno de difracción es muy débil en el caso de los cristales de proteína, el mayor brillo de la radiación sincrotrón permite obtener datos de mayor calidad. En cuanto a la sintonización de la radiación, posible en las instalaciones de radiación sincrotrón, es muy útil a la hora de resolver el problema de las fases necesarias para la resolución estructural.

3.3.3 Preparación de las muestras

Durante el experimento de difracción, la radiación de rayos X genera el calentamiento del cristal así como la aparición de radicales libres de las moléculas de proteína y de solvente, generando una degradación paulatina del cristal y la pérdida de orden interno. Para evitar al máximo estos daños irreversibles, la medida de los datos de difracción se lleva a cabo vitrificando la muestra bajo un flujo laminar de nitrógeno líquido a una temperatura de 100 K (Garman & Schneider, 1997).

Durante el proceso de vitrificación debemos evitar que el agua no estructural que hay en el interior del cristal de proteína forme hielo, ya que éste podría destruir nuestro cristal. Para ello, el cristal debe sumergirse en una solución crioprotectora y enfriarse rápidamente (McFerrin & Snell, 2002). La solución crioprotectora puede estar formada con sales, polímeros, alcoholes o azúcares, que reemplazan moléculas de agua y evitan la formación de cristales de hielo. El agente crioprotector utilizado debe escogerse con cuidado, ya que en algunos casos puede dañar la integridad del cristal. El cristal se sumerge durante unos segundos en la solución crioprotectora y a continuación se monta sobre un lazo de nylon inerte a los rayos X.

3.3.4 Procesado de los datos de difracción

Una vez se ha montado el cristal, éste es irradiado con rayos X mientras se le hace girar en intervalos adecuados, típicamente entre 0.5 y 1 grado. La medida realizada en los sucesivos intervalos nos proporcionan un conjunto de imágenes de difracción en la que se recoge la totalidad de los haces difractados por el cristal, estos haces constituyen el patrón de intensidades. En los puntos que se observan en las imágenes están contenidas las reflexiones. A partir del espectro de difracción vamos a poder conocer por un lado la celdilla unidad, que se calcula a partir de la posición relativa de los distintos haces difractados, y por otro lado el grupo espacial, que se obtiene del análisis de las posiciones de las intensidades.

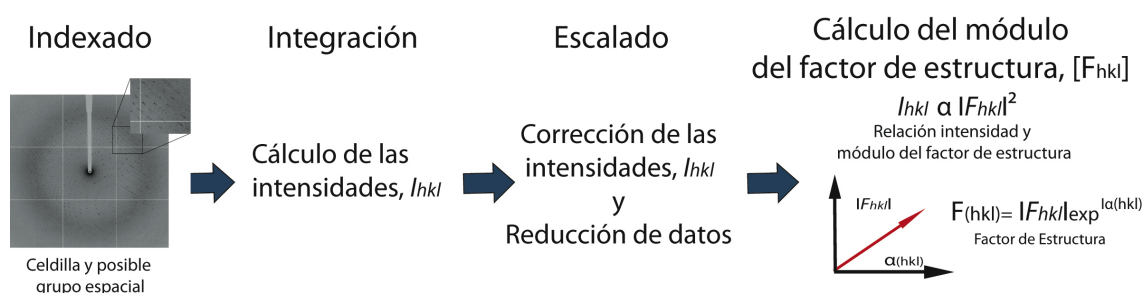


Figura 3.9. Esquema de los pasos necesarios para la estimación de las intensidades y del módulo del factor de estructura.

El procesado de los datos de difracción se divide en varias etapas, que se indican en el esquema que aparece a continuación:

Indexado: Se determinan los parámetros de la celdilla, calculándose sus dimensiones y topología. En esta etapa obtenemos estimación de los grupos espaciales más probables, seleccionando provisionalmente el que mejor se ajuste a las posiciones de las intensidades medidas. En esta etapa utilizamos el programa MOSFLM (Leslie, 1992).

Estimación de las intensidades: Durante la integración se lleva a cabo la medida de la intensidad de cada una de las reflexiones recogidas de las imágenes de difracción. En esta etapa se utiliza el programa MOSFLM (Leslie, 1992).

Escalado: Se lleva a cabo la corrección de intensidades teniendo en cuenta el ruido de fondo de las imagen, la mosaicidad del cristal, el decaimiento de la difracción del cristal y las fluctuaciones en la intensidad de la fuente de rayos X. A continuación se realiza un promediado de las intensidades llevando todos los datos a la misma escala relativa. Las intensidades ya corregidas y promediadas se reducen a la unidad asimétrica. Se utiliza el programa *SCALA* del paquete informático *CCP4* (Collaborative Computational Project, 1994). Algunos de los criterios de la calidad de los datos de difracción son:

R_{sym} : Indica las diferencias en la intensidad de las reflexiones relacionadas por simetría. Los valores globales para un buen conjunto de datos deben ser menores del 10%.

$I/\sigma(I)$: Representa la intensidad de la señal respecto del ruido. Son aceptables valores por encima de 2.

Multiplicidad o redundancia: Representa el número de veces que en promedio se ha medido cada reflexión. Cuanto mayor es su valor mejor será la estimación de los errores asociado a la medida de la reflexión.

Compleitud: Indica el porcentaje del espacio recíproco que ha sido medido. La completitud global necesaria comprende un 90-100%, así como del 80% en la última concha de resolución.

Cálculo del módulo del factor de estructura: El factor de estructura es la onda resultante del haz difractado por todos los átomos de la celdilla en una determinada dirección. Como cualquier onda, el factor de estructura viene caracterizado por su módulo, $|F_{hkl}|$ y su fase, α_{hkl} . El módulo representa la amplitud del haz difractado y puede calcularse directamente a partir de la intensidad medida para dicha reflexión. La fase, en cambio, se pierde durante el experimento de difracción. En esta etapa se utilizó el programa *TRUNCATE* del paquete informático *CCP4* (Collaborative Computational Project, 1994).

3.3.5 Resolución estructural

El fin último de la cristalografía de rayos X es el cálculo de la función de densidad electrónica, $\rho(xyz)$ (figura 3.10), que representa la envolvente electrónica de los átomos de la molécula dentro del cristal. Haciendo uso de la teoría de interferencias de ondas es fácil demostrar que la geometría del espectro de difracción da cuenta de la celdilla unidad, y la simetría de las intensidades informa de la simetría cristalina. Además, se puede establecer una relación directa entre la estructura cristalina y el experimento mediante la transformada de Fourier (figura 3.10).

$$\rho(xyz) = 1/V \sum_h \sum_k \sum_l |F(hkl)| \cos 2\pi (hx + ky + lz - \alpha(hkl))$$

Amplitudes Fases

Figura 3.10. Función de densidad electrónica. V es el volumen de la celdilla unidad y $\alpha(hkl)$ es la fase asociada al módulo del factor de estructura, $|F(hkl)|$. Es interesante resaltar el hecho de que la densidad electrónica en un punto viene determinada por la información de todo el espectro. Inversamente, el valor de un factor de estructura (módulo y fase) viene determinado por las posiciones de todos los átomos que forman la estructura, $\rho(xyz)$.

Esta expresión relaciona la densidad electrónica, $\rho(xyz)$, en cada punto (xyz) del cristal, es decir la estructura, y el experimento de difracción, donde $|F(hkl)|$ es proporcional a la raíz cuadrada de la intensidad de cada punto (hkl) del espectro de difracción. Para poder resolver la ecuación es necesario disponer no sólo de los módulos, $|F(hkl)|$, sino también de las fases, $\alpha(hkl)$, asociadas a los factores de estructura. Esta información, sin embargo, se pierde durante el experimento. Esto constituye el denominado “problema de la fase” y el trabajo del cristalógrafo consiste en averiguar el valor de $\alpha(hkl)$ mediante diferentes metodologías.

En el caso de las moléculas de pequeño y medio tamaño el problema de la fase puede resolverse mediante los llamados métodos directos. En el caso de las proteínas, su gran tamaño genera un elevado número de variables y no permite la utilización de estos métodos directos y hemos de recurrir a los llamados métodos indirectos. En cristalografía de macromoléculas existen principalmente tres métodos para calcular las fases: (i) empleando el método del **Reemplazo Molecular (MR)**, del inglés *Molecular Replacement*), haciendo uso de un modelo estructural de una proteína homóloga, previamente determinada, (ii) introduciendo átomos altamente dispersores, o método del **Reemplazo Isomorfo Múltiple (MIR)**, del inglés, *Multiple Isomorphous Replacement*) o (iii) introduciendo átomos dispersores anómalos, o método de **Dispersión Anómala Simple o Múltiple (SAD)**, del inglés *Single-wavelength Anomalous Diffraction* o **MAD**, del inglés *Multi-wavelength Anomalous Diffraction*). Con cualquiera de estos tres métodos podemos calcular unas fases iniciales, $\alpha_c(hkl)$, que, junto con las amplitudes experimentales, $|F_{\text{exp}}(hkl)|$, nos permitirán calcular una función de densidad electrónica aproximada sobre la cual podremos construir nuestro modelo estructural.

Una vez que tenemos la cadena polipeptídica trazada, tenemos un modelo inicial, aún insuficiente para definir la estructura de nuestra proteína y que se mejora mediante sucesivas etapas refinamiento matemático y modelado manual.

REFINAMIENTO MATEMÁTICO

El Refinamiento permite ajustar los parámetros atómicos que definen nuestro modelo estructural. Estos son:

Coordenadas atómicas (x,y,z): que determinan la posición de cada átomo dentro de la celdilla unidad.

Factor térmico B: Representa la vibración térmica o movilidad de cada átomo.

Ocupación: Hace referencia a la fracción de moléculas en las que un átomo ocupa una determinada posición.

Dado el elevado tamaño de las moléculas de proteína, el número de parámetros a refinar es del orden al número de observaciones, por lo que es necesario introducir una serie de restricciones geométricas (distancias y ángulos de enlace, torsiones...), que en el caso de los aminoácidos y otras moléculas son conocidas con gran precisión.

El método de refinamiento utilizado depende del estado del modelo. En las primeras etapas del refinamiento se aplican los métodos de refinamiento de Cuerpo Rígido y Calentamiento Simulado (*Simulated Annealing*), mientras que en etapas posteriores se utilizan los Métodos Probabilísticos de Máxima Verosimilitud (*Maximum Likelihood*).

Durante el proceso de refinamiento debemos asegurarnos que los cambios introducidos en el modelo converjan cada vez más hacia la estructura real de nuestra molécula. Para seguir esta convergencia calculamos el factor de desacuerdo R , que nos indica la bondad del modelo mediante el cálculo de la diferencias entre el espectro observado ($|F_o|$) y el calculado a partir de el modelo experimental ($|F_c|$).

El factor R debe ir disminuyendo a lo largo del refinamiento, pero un valor bajo no justifica que el modelo sea correcto. Para asegurarnos de que los cambios que realizamos sobre el modelo son representativos de la estructura real utilizamos un Método Estadístico de Validación Cruzada basado en el cálculo de un segundo factor de convergencia, el R_{free} . Este método implica la separación del total de las reflexiones en dos conjuntos, un 90-95% de las reflexiones que serán utilizadas en el refinamiento (*working test*) y un 5-10% de las reflexiones que no se usarán en el refinamiento (*test set*). El factor R_{free} va a medir el desacuerdo que existe entre el modelo refinado y las reflexiones que no utilizamos en el refinamiento, por lo que sólo los cambios que disminuyan el R_{free} serán correctos.

$$R = \frac{\sum_{hkl} |F_{obs}| - k |F_{cal}|}{\sum_{hkl} |F_{obs}|} \times 100\%$$

Para el Refinamiento Matemático se emplearon los programas como *CNS* (Brunger *et al.*, 1998), *REFMAC* (Murshudov *et al.*, 1997) y *PHENIX* (Adams *et al.*, 2010).

MODELADO MANUAL Y VALIDACIÓN DEL MODELO FINAL

Algunos de los errores del modelo no pueden solventarse mediante refinamiento matemático y han de corregirse manualmente en estaciones gráficas. Estos errores pueden deberse a:

- Errores en la colocación de la cadena principal o lateral de los aminoácidos.

- Errores de secuencia. Diferencias entre modelo de Reemplazo Molecular y nuestra molécula.
- Errores en la conectividad de los átomos. Existencia de inserciones o eliminaciones de aminoácidos entre el modelo de Reemplazo Molecular y la estructura de nuestra molécula.
- Presencia de ligandos ausentes en el modelo inicial.

La corrección de estos errores se realiza sobre los mapas de densidad electrónica calculados mediante una síntesis de Fourier, con los factores de estructura observados (F_o) y los factores de estructura calculados para el modelo (F_c). Los coeficientes utilizados en el cálculo de los mapas fueron $2F_o - F_c$ y $F_o - F_c$.

El refinamiento se da por finalizado cuando el modelo se ajusta al mapa de densidad electrónica y los factores R y R_{free} , se encuentran en unos valores cercanos al 20% (R) y menores al 30% (R_{free}). Llegado este punto debemos comprobar que los parámetros tales como las distancias y ángulos de enlace, la conformación de los rotámeros y los ángulos de torsión sean razonables. El diagrama de Ramachandran presenta los ángulos de torsión Ψ y Φ de los residuos de la cadena polipeptídica, de esta manera podemos conocer si los enlaces entre los aminoácidos se encuentran dentro del rango permitido. Estos parámetros son analizados con el programa *PROCHECK* (Laskowski *et al.*, 1993) y *COOT*, (Emsley & Cowtan, 2004).

3.4 ENSAYOS DE “DOCKING” CON HERRAMIENTAS INFORMÁTICAS

3.4.1 Programa *GOLD*

GOLD (de sus siglas en inglés *Genetic Optimization for Ligand Docking*) (Verdonk *et al.*, 2003), es un programa para calcular modelos de “docking” entre pequeñas moléculas (ligandos) en posibles sitios de unión de una proteína, basado en un algoritmo genético. Nos permite asignar la cavidad de unión del ligando, y busca aquellas orientaciones en las que el ligando puede unirse en el interior de dicha cavidad, tomando en cuenta las interacciones posibles. Por defecto, *GOLD* trata al ligando como una molécula flexible, mientras que la proteína, es considerada como un cuerpo rígido. La flexibilidad del ligando puede modificarse introduciendo constricciones en sus ángulos, distancias de enlace, etc. En el caso de la proteína, *GOLD* permite asignar flexibilidad en las cadenas laterales de hasta un máximo de 10 residuos. La flexibilidad de la cadena lateral de un residuo se determina introduciendo el valor inicial del ángulo y el intervalo de cada uno de los ángulos Chi de la cadena lateral que queramos ensayar. El número de soluciones se suele limitar a 10, aunque cuando el programa encuentra tres soluciones con los mismos mínimos de energía se detiene.

GOLD requiere la presencia de los átomos de hidrógeno en los modelos estructurales y trabaja con ficheros de coordenadas con formato “mol2”. Para introducir los átomos de hidrógeno en los ficheros de coordenadas utilizados y reeditarlos en formato “mol2” se utilizó el programa *SYBYL* (Tripos Inc., Louis. Mo Tripos, USA). El servidor “*The Dundee PRODRG*”, fue utilizado en los casos de

no disponer información estructural, el cual nos permite obtener las coordenadas atómicas de moléculas pequeñas, pasando de un esquema en 2D a 3D, aplicando bases de datos estructurales de moléculas pequeñas y restricciones geométricas (<http://davapc1.bioch.dundee.ac.uk/prodrg>). De igual forma, para construir las moléculas de carbohidratos de las cuales no se dispone de información estructural se utilizó la aplicación “sweet2” del servidor “glycosciences”, el cual nos permite modelar estructuras de largas cadenas polisacarídicas (www.glycosciences.de/index.php).

3.5 OTROS PROGRAMAS UTILIZADOS

Alineamiento de secuencias: programa BLAST (Altschul et al., 1990) <http://www.expasy.ch/tools/blast/> y se el programa CLUSTALW (Thompson et al., 1994) <http://www.ebi.ac.uk/Tools/clustalw2/>.

Análisis de la estructura secundaria: Se utilizaron los programas JOY (Overington et al., 1990; Mizuguchi et al., 1998) y PROCHECK (Laskowski et al., 1993).

Representación gráfica y elaboración de figuras: Se utilizó el programas PyMOL (DeLano 2002) www.pymol.org.

Servidor Dalí: Se utiliza para buscar similitud entre estructuras tridimensionales de proteínas utilizando la base de datos de PDB (de sus siglas en inglés: protein data bank). El resultado puede mostrar similitudes biológicamente interesantes que no son detectables mediante la comparación de secuencias (Holm, L. & Rosenström, P. 2010) http://ekhidna.biocenter.helsinki.fi/dali_server/start.

3.6 TÉCNICAS COMPLEMENTARIAS UTILIZADAS

3.6.1 Espectrometría de masas

La espectrometría de masas es una técnica basada en la ionización de las moléculas y en su separación en función de la relación masa/carga (m/z).

Los experimentos se realizaron en el Servicio de Espectrometría de Masas del Instituto de Química-Física Rocasolano (CSIC) y se llevaron a cabo en un espectrómetro de masas Voyager DE-PRO (Applied Biosystems) equipado con un láser de nitrógeno ($\lambda = 337$ nm; anchura de pulso = 3 ns; $\nu = 20$ Hz) y una fuente iónica con extracción retardada. Los iones positivos generados por la resonancia láser fueron introducidos en el tubo de vuelo (1.3 m de longitud) con un voltaje de aceleración de 25 kV, trabajando en modo lineal. Todos los espectros de masas se obtuvieron recogiendo una media de 5000 disparos, y la matriz utilizada fue ácido sinapínico ($10 \text{ mg} \cdot \text{ml}^{-1}$ en una mezcla TFA al 0.21% y acetonitrilo al 30%). La calibración necesaria para los espectros de masas fue externa, empleándose apomioglobina de caballo.

3.6.2 Secuenciación de proteínas

La obtención de la secuencia de aminoácidos del extremo N-terminal fueron llevados a cabo en el servicio de química de proteínas del Centro de Investigaciones Biológicas (CSIC) en un

secuenciador automático modelo “Applied Biosystems” (Perkin-Elmer Procise 494). Permite determinar la secuencia de péptidos y proteínas mediante degradación secuencial de Edman (Pehr, 1950). El reactivo de Edman o fenilisotiocianato, se utiliza en condiciones ligeramente básicas, para marcar específicamente el grupo amino terminal de una proteína. El derivado feniltioidantoina (comúnmente conocido como PTH), hace inestable el segmento N-terminal de manera que puede ser selectivamente hidrolizado, sin romper los demás enlaces peptídicos. El péptido acortado resultante es analizado por cromatografía líquida de alta presión, este patrón se compara con el original y se puede determinar el aminoácido faltante (N-terminal), por la diferencia del tiempo de retardo en la cromatografía. El reactivo de Edman puede ser aplicado repetidamente al péptido resultante del tratamiento anterior y de esta forma se pueden llegar a determinar la secuencia de los primeros 100 aminoácidos. Se requieren aproximadamente 50-100 pmol de muestra pura (proteína/péptido), que puede estar en solución o electrotransferidas a una membrana polivinilo.

3.7 RESOLUCIÓN ESTRUCTURAL DE LA ENDOLISINA Cpl-7

3.7.1 Obtención de las muestras

- Producción de la proteína Cpl-7 completa

El gen que codifica la proteína Cpl-7 madura, fue clonado dentro del plásmido pCP700, (García *et al.*, 1990). El fragmento de la PCR correspondiente fue amplificado con los cebadores (5'-**GGAATTC**ATGGTTAAGAAAAATGATTTATTTGTAGACGTTG-3') en sentido y el (5'-**CCCAAGCTT**AAAATAGCTAGTAGAAAATTTCTACTAGCTTTTACTTGTTA-3') antisentido, en donde los nucleótidos en negrita representan los sitios de corte con las enzimas de restricción EcoR1 y HindIII (TAKARA), respectivamente. El fragmento resultante de la PCR que contenía un peso de 1000 kb fue purificado y subclonado dentro del vector de expresión pKLSL_t, el cual codifica una proteína de fusión LSL_t-Cpl-7. Con esta nueva construcción se transformó la cepa competente BL21 (DE3) de *E. coli* (Promega), para su expresión y purificación. Esta construcción contiene el gen que codifica para la proteína Cpl-7 madura con 344 aminoácidos, de peso molecular 38 kDa y un coeficiente de extinción molar teórico de 68425 M⁻¹·cm⁻¹ a 280 nm.

El vector pKLSL_t expresa la proteína de interés con una proteína de fusión en su extremo N-terminal para facilitar su purificación mediante cromatografía de afinidad. El vector pKLSL_t ha sido construido sobre el plásmido pET28a(+) (Novagen, Germany), al cual se le ha incorporado una secuencia de nucleótidos que codifica para la proteína LSL_t, perteneciente al módulo lectina de la proteína hemolítica LSL_a del hongo *Laetiporus sulphureus* (Mancheño *et al.*, 2005), seguido por una secuencia que codifica para un conector corto (ASSS) y un sitio de corte para la endopeptidasa TEV (de sus siglas en inglés: *tobacco etch virus*) del virus del mosaico del tabaco (ENLYFQG) (Mancheño & Angulo, 2009).

Las bacterias se crecieron en medio de LB (Luria-Bertani: 10 g·l⁻¹ triptona, 5 g·l⁻¹ extracto de levadura, 10 g·l⁻¹ NaCl, pH 7.5). Para el crecimiento de bacterias transformadas con plásmidos, el

medio se suplementó con el antibiótico kanamicina ($50 \mu\text{g}\cdot\text{ml}^{-1}$) (Sigma). Los cultivos de colonias en medio sólido se realizaron en el mismo medio con 1.5% bacto-agar (Difco) y antibióticos a la concentración indicada anteriormente.

Para la purificación de la forma completa, Cpl-7 fue expresada en la cepa BL21 crecidas en medio LB (Luria-Bertani) a 37°C . Cuando el cultivo alcanzó la fase de crecimiento logarítmico a una densidad óptica aproximada de 0.8 (longitud de onda de 600 nm) se añadió $50 \mu\text{M}$ de isopropil- β -D-tiogalactopiranosido (IPTG), para activar la expresión de la proteína de fusión y la incubación se continuó a 16°C durante 12 horas. Las células se recogieron en la fase estacionaria de crecimiento mediante centrifugación a $10000 \times g$ durante 20 minutos, resuspendiéndose posteriormente en tampón A (Tris-HCl 20 mM, pH 8 y NaCl 0.1 M), suplementado con un cóctel de inhibidores de proteasa (ROCHE). A continuación se lisaron las células utilizando una prensa french (1100 psi) y el extracto celular se centrifugó a $17000 \times g$ durante 40 minutos. La proteína Cpl-7 se purificó a partir de la fracción soluble siguiendo el procedimiento descrito a continuación:

1. Cromatografía de afinidad en columna Sepharosa 4B (SIGMA), seguido por un lavado de la columna con tampón A, para eliminar impurezas retenidas en la columna de forma inespecífica, a una velocidad de 1 ml min^{-1} y las fracciones colectadas de 4 ml.
2. Separación de la resina con tampón A suplementado con lactosa 200 mM.
3. Incubación de la proteína de fusión LSLt-Cpl-7 con la proteasa TEV (proporción 80:1), para separar ambas proteínas.
4. Cromatografía de exclusión molecular en columna Superdex 75 en sistema FPLC (BioRad), en tampón A. Para separar las proteínas por su tamaño molecular a una velocidad de $0.5 \text{ ml}\cdot\text{min}^{-1}$ y las fracciones colectadas de 2 ml.

- Producción del dominio C-terminal de Cpl-7

El dominio C-terminal de Cpl-7 fue clonado a partir del plásmido pCP700 el cual contenía el gen que codifica a la proteína Cpl-7 completa, utilizando como cebadores, (5'-**GGAATTCA**ACAATGAAAACACTCTAAAAAGCCTTACC-3') en sentido y (5'-CCCA**AAGCTT**AAAATAGCTAGTAGAAAATTTCTACTAGCTTTTACTTGTTA-3') antisentido (SIGMA), en donde los nucleótidos en negrita representan los sitios de corte con las enzimas de restricción EcoR1 y HindIII (TAKARA), respectivamente. El fragmento resultante de la PCR que contenía un peso de 500 kb fue purificado y subclonado dentro del vector de expresión pKLSL_t el cual resulta en una proteína de fusión LSL_t-CCpl-7. Con esta nueva construcción se transformó la cepa competente BL21 (DE3) de *E. coli* (Promega), para su expresión y purificación siguiendo el mismo protocolo descrito para Cpl-7 completa. Esta construcción contiene el gen que codifica para el módulo C-terminal de la proteína Cpl-7 con 142 aminoácidos, de peso molecular 16 kDa y un coeficiente de extinción molar teórico de $25440 \text{ M}^{-1}\cdot\text{cm}^{-1}$ a 280 nm.

Todo el proceso se llevó a cabo a 4 °C, se registraron las distintas fracciones de separación mediante espectros de absorción a una longitud de onda de 280 nm comprobando de este modo la purificación. La homogeneidad de las muestras se verificó mediante electroforesis en geles desnaturalizantes de poliacrilamida (SDS-PAGE) al 12% siguiendo el método de Laemmli (Laemmli, 1970) y espectroscopia de masas. Las proteínas purificadas fueron concentradas utilizando un concentrador Amicon YM-10 con un corte de 10 kDa. Para estimar la concentración final de las muestras se midió la absorbancia en un espectrofotómetro Nanodrop ND-1000 (Thermo Scientific), usando los coeficientes de extinción molar teóricos a 280 nm.

3.7.2 Cristalización de Cpl-7

- Cristalización de Cpl-7 completa

Los ensayos iniciales de cristalización se llevaron a cabo utilizando técnicas de alto rendimiento con el método de difusión de vapor en gota sentada a 291 K (ver sección 3.3.1). Se mezclaron 250 nl de Cpl-7 purificada ($10 \text{ mg} \cdot \text{ml}^{-1}$ en tampón Tris-HCl 20 mM, pH 8 y NaCl 0.1 M) y 250 nl de solución precipitante, equilibradas con 80 μl de reservorio, ensayando alrededor de 1300 condiciones. Luego de observaciones sucesivas en el tiempo, se escogieron aquellas condiciones en donde se observó la aparición de pequeños cristales y precipitados cristalinos, para el proceso de optimización.

- Cristalización del módulo de anclaje de Cpl-7

Las condiciones iniciales de cristalización se establecieron empleando técnicas de alto rendimiento con el método de difusión de vapor en gota sentada a 291 K (ver sección 3.3.1), mezclando 250 nl o 500 nl de CCpl-7 ($10 \text{ mg} \cdot \text{ml}^{-1}$, $20 \text{ mg} \cdot \text{ml}^{-1}$ y $40 \text{ mg} \cdot \text{ml}^{-1}$ en tampón Tris-HCl 20 mM, pH 8 y NaCl 0.1 M). Al cabo de un tiempo crecieron cristales de pequeño tamaño en diversas condiciones. A partir de estas condiciones iniciales de cristalización se realizaron sucesivos ensayos por el método de difusión de vapor en gota colgante (ver sección 3.3.1), mezclando 1 μl de la solución de proteína con 1 μl de la solución de cristalización y las gotas se equilibraron frente a 500 μl de reservorio. Haciendo pequeñas variaciones de las condiciones iniciales, añadiendo aditivos y detergentes (Hampton Research, CA), para mejorar el tamaño y calidad de los cristales.

3.7.3 Resolución estructural de Cpl-7

- Resolución estructural de Cpl-7 completa

Se recogieron datos de difracción por medio de un detector CCD Kappa 2000 Brüker-Nonius empleando la radiación $\text{CuK}\alpha$ ($\lambda = 1.5418 \text{ \AA}$) proporcionada por una fuente de ánodo rotatorio Brüker-Nonius FR591, equipado con un sistema de focalización de doble espejo y operando a 45 kV y 100 mA. La distancia del cristal al detector fue 200 mm. Los cristales fueron crioprotectados por inmersión en una solución que contenía 30% de glicerol en la solución madre y se vitrificaron rápidamente a 100 K en una corriente de nitrógeno generada por el sistema criogénico (Oxford Cryosystems).

Las imágenes de difracción fueron procesadas empleando *DENZO* y *SCALEPACK* pertenecientes al paquete de programas *HKL-2000*, (Otwinowski *et al.*, 2003). Las intensidades fueron convertidas en factores de estructura usando el programa *TRUNCATE* perteneciente al paquete de programas de *CCP4* (Collaborative Computational Project, 1994). La estructura fue resuelta mediante reemplazo molecular utilizando como modelo inicial el módulo catalítico de Cpl-1 (código PDB: 2IXU) empleando el programa *MOLREP* (Collaborative Computational Project, 1994; Vagin & Teplyakov, 1997), el modelo fue refinado con el programa *CNS* (Brunger *et al.*, 1998), el ajuste y mutaciones de los aminoácidos en el mapa de densidad fue realizado manualmente utilizando el programa *O* (Jones *et al.*, 1991) y la validación de la estructura se realizó con el programa *PROCHECK* (Collaborative Computational Project 1994).

- Resolución estructural del módulo de anclaje de Cpl-7

Los cristales optimizados del módulo C-terminal de Cpl-7 se vitrificaron en nitrógeno líquido directamente tras recogerlos de las gotas. Los datos fueron medidos en una fuente de radiación sincrotrón en la línea ID23-1 del ESRF (*European Synchrotron Radiation Facility*), Grenoble – Francia. A una longitud de onda de 0.9790 Å a 100 K. El detector usado para la recolección de los patrones de difracción fue un ADSC Quantum Q315r.

Las imágenes de difracción fueron procesadas y escaladas empleando el programa *iMOSFLM* y *SCALA* del paquete informático *CCP4*. La estructura fue resuelta utilizando el programa *ARCIMBOLDO* (Rodríguez *et al.*, 2009). Este programa se basa en la búsqueda de pequeños fragmentos de hélices- α de poli-Ala con el programa *PHASER* (McCoy *et al.*, 2007), con las mejores soluciones se realiza una modificación de la densidad electrónica con el programa *SHELX/E* (Sheldrick, 2008). Una vez mejorados los mapas de densidad electrónica se trazó el modelo con la secuencia de aminoácidos de la proteína usando el programa gráfico *COOT* (Emsley & Cowtan, 2004). El modelo inicial fue sometido a sucesivos ciclos de refinamiento y colocación de moléculas de agua. La validación de la estructura se realizó con el programa *PHENIX*, (Adams *et al.*, 2010).

3.8 RESOLUCIÓN ESTRUCTURAL DE SIGN-R1

3.8.1 Obtención de las muestras

La forma soluble, tanto del dominio de unión a carbohidratos (CRD) como del dominio extracelular (ECD) de SIGN-R1, fueron producidas con un péptido señal para ser secretadas al medio y una etiqueta para su purificación en el extremo N-terminal. Así, estas construcciones fueron utilizadas para producir la proteína en sistema de mamíferos, transfectando células del ovario del hámster chino (siglas en inglés *CHO cell*). La proteína fue purificada a partir del sobrenadante del cultivo con una columna de afinidad cargada con anticuerpos que reconocen la secuencia de aminoácidos de la etiqueta (DYKDDDDKL) (Sigma Aldrich). Se obtuvo 1 mg de proteína que se concentró a $3.5 \text{ mg}^{-1} \cdot \text{ml}^{-1}$, centrifugando a través de un concentrador Amicon, cuya membrana tenía un corte de 10 kDa en tampón Tris-HCl 20 mM, pH 7.5 y NaCl 0.1 M. Esta proteína fue producida por

el grupo del Dr. Chae Gyu Park en el “Laboratory of Cellular Physiology and Immunology in the Chris Browne Center for Immunology and Immune Diseases, The Rockefeller University” NY. La proteína madura del dominio extracelular de SIGN-R1 (ECD_SIGN-R1), tiene 259 aminoácidos y un peso molecular 30 kDa, mientras que la construcción truncada del módulo de unión a carbohidratos de SIGN-R1 (CRD_SIGN-R1), tiene 143 aminoácidos y un peso molecular de 16 kDa, y un coeficiente de extinción molar teórico de $5546 \text{ M}^{-1} \cdot \text{cm}^{-1}$ y $53970 \text{ M}^{-1} \cdot \text{cm}^{-1}$ a 280 nm, respectivamente.

3.8.2 Cristalización del ECD_SIGN-R1 y CRD_SIGN-R1

Las condiciones iniciales de cristalización se establecieron empleando técnicas de alto rendimiento con el método de difusión de vapor en gota sentada a 291 K (ver sección 3.2.3), utilizando la batería de soluciones comerciales PACT suite y JCSG suite (Quiagen) y el JBScreen Classics 1,4,5 y 7 (Jena Bioscience), mezclando 250 nL de CRD_SIGN-R1 a $3.5 \text{ mg} \cdot \text{mL}^{-1}$ en Tris-HCl 20 mM, pH 7.5 y NaCl 0.1 M. Al cabo de un tiempo crecieron cristales de pequeño tamaño en diversas condiciones. A partir de estas condiciones iniciales de cristalización se realizaron sucesivos ensayos por el método de difusión de vapor en gota colgante (ver sección 3.2.3), mezclando 1 μL de la solución de proteína con 1 μL de la solución de cristalización y las gotas se equilibraron frente a 500 μL de reservorio. Haciendo pequeñas variaciones de las condiciones iniciales, se consiguió mejorar la calidad y tamaño de los cristales. Para el caso de ECD_SIGN-R1 se ensayó un gran número de condiciones, aditivos y detergentes y no se obtuvieron cristales.

3.8.3 Cristalización de complejos CRD_SIGN-R1 : ligando

Para la búsqueda y preparación de los complejos de la proteína CRD_SIGN-R1 : ligando, se llevaron a cabo gran cantidad de ensayos con una amplia batería de carbohidratos. Los procesos se llevaron a cabo tanto por inmersión, como por cocrystalización (ver sección 3.3.4.). Las soluciones en todos los casos estaban suplementadas con cloruro de calcio 5 mM. Los ligandos ensayados, el método y resolución estructural se detallan en la siguiente tabla:

| Ligando | [concentración] | Método | Medida de difracción ESRF | Observación del ligando |
|--------------------------------------|-----------------|---------------------|---------------------------|-------------------------|
| D-Manosa-6-Fosfato | 25 mM | Inmersión 10' y 24h | ID14-2 | - |
| D-Galactosa-6-Fosfato | 50 mM | Inmersión 10' y 24h | ID14-2 / BM_16 | - |
| N-Acetil-Galactosamina | 50 mM | Inmersión 10' y 24h | ID14-2 | - |
| Metil-D-Manopiranososa | 50 mM | Inmersión 10' y 24h | ID14-2 / BM_16 | - |
| Fosforilcolina | 50 mM | Inmersión 10' y 24h | ID14-2 | - |
| Mezcla: Pcho+ N-Acetil-Galactosamina | 50 mM : 50mM | Inmersión 10' y 24h | ID14-2 | - |
| Calcio | 10 mM | Inmersión 24h | ID14-2 | - |

| | | | | |
|--------------------|---------------------------|--|---------------|----------|
| | | | | |
| Man-9 | 1 mM | Inmersión 12h Cocrystalización | ID14-3 | - |
| Man-3 | 1 mM | Inmersión 12h Cocrystalización | ID14-3 | - |
| Á. Siálico | 40 mM | Inmersión 10' Cocrystalización | ID14-3 | + |
| Lewis X | 10 mM | Inmersión 10' Cocrystalización | ID14-3 | - |
| S. Dextrano | 15 % 3 % | Inmersión 10' y 3h Cocrystalización | ID14-3 | + |
| Á. Teicoico | 5,10,15 y 25 mM | Inmersión 10' y 24h Cocrystalización | ID14-3 | - |
| Di-Manosa | 50 mM | Inmersión 30" y 3h | ID14-1 | - |
| Tetra-Manosa | 50 mM | Inmersión 30" y 3h | ID14-1 | - |

Figura 3.11. Ensayos experimentales para la obtención de los complejos CRD_SIGN-R1 : Ligando. H (horas); - (no se observó el ligando); + (se observó el ligando).

Todos los compuestos reflejados en la figura 3.11 se adquirieron de Sigma Aldrich, excepto el Sulfo-dextrano que se obtuvo de "Hampton Research". Los ligandos Di-Manosa y Tetra-Manosa fueron sintetizados en CIC bioGUNE (*Center for Cooperative Research in Biosciences*) en grupo de la Dra. Soledad Penades y los fragmentos de peptidoglicano han sido purificados en el laboratorio de Waldemar Vollmer (Universidad de Newcastle - Reino Unido).

3.8.4 Determinación estructural de CRD_SIGN-R1 y sus complejos

Todos los cristales fueron vitrificados en nitrógeno líquido, previamente crio-protectados por inmersión en una solución que contenía 50% de sulfato de litio saturado en la solución madre. Los datos de difracción para CRD_SIGN-R1 nativa fueron medidos en la línea ID23-1 del ESRF (*European Synchrotron Radiation Facility*), Grenoble – Francia. A una longitud de onda de 1.072250 Å a 100 K. El detector usado para la recolección de los patrones de difracción fue un ADSC Quantum Q315r. Las imágenes de difracción fueron procesadas y escaladas empleando el programa *iMOSFLM* y *SCALA* del paquete informático *CCP4*. La estructura fue resuelta por remplazo molecular con el programa *MOLREP* (Vagin & Teplyakov, 1997), utilizando como modelo la estructura de su homólogo en células dendríticas DC_SIGN (código PDB: 1K9I), la cual tiene una identidad de 68%. Una vez obtenidas las fases, se refinó el modelo inicial y se calcularon los mapas de densidad electrónica usando el programa *CNS*, la cadena principal y las cadenas laterales fueron ajustadas usando el programa *O* (Brunger *et al.*, 1998). El modelo fue sometido a sucesivos ciclos de refinamiento, colocación de moléculas de agua y la validación de la estructura se realizó con el programa *PHENIX* (Adams *et al.*, 2010).

El complejo de la molécula de sulfo-dextrano con el CRD_SIGN-R1 fue obtenido por cocrystalización, mezclando la solución de proteína con el ligando durante 4 horas antes de montar los ensayos de cristalización por el método de gota colgante. Los cristales fueron medidos en la línea ID14-1 del ESRF. A una longitud de onda de 0.93340 Å a 100 K. El detector usado para la recolección de los patrones de difracción fue un ADSC Quantum 4. Mientras que los cristales del complejo con ácido siálico fueron obtenidos por inmersión y los datos de difracción fueron recogidos en un detector MAR225 de la línea ID23-2 del ESRF.

Las imágenes de difracción de los complejos fueron procesadas y escaladas empleando el programa *iMOSFLM* y *SCALA* del paquete informático *CCP4*. Las estructuras fueron resueltas por remplazo molecular con el programa *MOLREP*, (Jones *et al.*, 1991; Vagin & Teplyakov, 1997) utilizando como modelo la estructura del CRD_SIGN-R1 nativo. Una vez obtenidas las fases, se refinó el modelo inicial y se calcularon los mapas de densidad electrónica usando el programa *CNS*, la cadena principal y las cadenas laterales fueron ajustadas usando el programa *O* (Jones T. A. *et al.* 1991) y *COOT* (Emsley & Cowtan, 2004). La molécula de sulfo-dextrano fue construida en el servidor “The Dundee PRODRG” y la molécula de ácido siálico fue tomada de las librerías de *CCP4*. El modelo fue sometido a sucesivos ciclos de refinamiento, colocación de moléculas de agua y la validación de la estructura se realizó con el programa *PHENIX* (Adams *et al.*, 2010).

3.8.5 Ensayos de “docking” CRD_SIGN-R1 con ribitol fosfato y Cápsula polisacárida

Se realizaron ensayos de “docking” entre el CRD_SIGN-R1 y la molécula de 5-ribitol-fosfato y un fragmento de la cápsula polisacárida del *S. pneumoniae* del serotipo 14 (figura 3.12), utilizando el programa *GOLD* y procedimientos estándar para la flexibilidad molecular con *ICM*, respectivamente. La molécula fue construida en el servidor “The Dundee PRODRG”. Debido a la forma alargada del ribitol-fosfato se realizaron algunas restricciones de distancia entre los átomos del ligando y los átomos de la proteína, así como también se le asignaron grados de libertad a los aminoácidos expuestos al solvente como lisinas y argininas. De entre todas las posibles soluciones se seleccionaron las de menor energía libre y mayor puntuación según el programa *GOLD*.

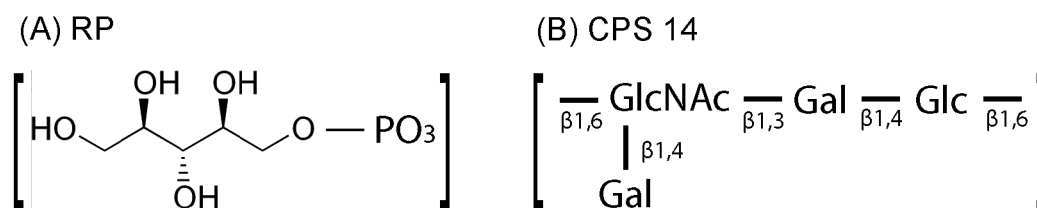


Figura 3.12. Representación esquemática de los ligandos utilizados en los ensayos de “docking”. (A) Ribosa-5-fosfato, unidad repetitiva que daría lugar a una cadena de poli-ribitol-fosfato enlazados por enlaces fosfodiéster. (B) Unidad repetitiva de la cápsula polisacárida de *S. pneumoniae* del serotipo 14; Glc, glucosa; Gal, galactosa; GlcNAc, N-acetil-glucosamina; β#.#, indica el tipo de enlace glicosídico y los carbonos por donde son enlazados.

3.9 DETERMINACIÓN ESTRUCTURAL DE LA PORCIÓN Fc DE IgG

3.9.1 Obtención de las muestras

La inmunoglobulina G (IgG) se produjo de forma recombinante y subsecuentemente glicosiladas *in vitro*. Para añadir el ácido siálico en la posición final de la cadena glicánica se utilizó la enzima 2,6-sialil-transferasa y luego estos anticuerpos fueron tratados con la proteasa papaína para separarlo en dos fragmentos, el Fab y el Fc. El fragmento Fc sialilado (Fc_Sial), fue proporcionado por el grupo de el Dr. Robert M. Anthony del Laboratorio de Genética Molecular e Inmunología de la Universidad Rockefeller NY. Cuyo proceso de purificación y glicosidación está descrito en (Anthony *et al.*, 2008a; Anthony *et al.*, 2008b). Una cantidad de 8.2 mg de Fc_Sial pura y homogénea se encontraba en tampón fosfato salino (PBS: Na₂HPO₄ 10 mM, KH₂PO₄ 1.76 mM, pH 7.5, NaCl 37 mM y KCl 2.5 mM) a una concentración de 9.5 mg·ml⁻¹. El Fc_Sial tiene 227 aminoácidos, un peso molecular de 25.5 kDa y un coeficiente de extinción molar teórico de 35785 M⁻¹·cm⁻¹ a 280 nm.

3.9.2 Cristalización

Las condiciones iniciales de cristalización se establecieron empleando técnicas de alto rendimiento con el método de difusión de vapor en gota sentada a 291 K (ver sección 3.2.3), utilizando la batería de soluciones comerciales PACT suite y JCSG suite (Quiagen) y el JBScreen Classics 1,4,5 y 7 (Jena Bioscience), mezclando 250 nl de Fc_Sial a 9.5 mg·ml⁻¹ en BPS. Al cabo de poco tiempo crecieron microcristales y agujas de pequeño tamaño en diversas condiciones. A partir de estas condiciones iniciales de cristalización se realizaron sucesivos ensayos por el método de difusión de vapor en gota colgante (ver sección 3.2.3), mezclando 1 µl de la solución de proteína con 1 µl de la solución de cristalización y las gotas se equilibraron frente a 500 µl de reservorio. Haciendo pequeñas variaciones de las condiciones iniciales así como también variaciones en la porción de la gota, se consiguieron cristales de buena calidad y tamaño.

3.9.3 Resolución estructural

Los cristales fueron vitrificados en nitrógeno líquido, previamente crio-protectados por inmersión en una solución que contenía glicerol. Los datos de difracción fueron medidos en la línea ID14-4 del ESRF a una longitud de onda de 0.99657 Å y a 100 K. El detector usado para la recolección de los patrones de difracción fue un ADSC Quantum Q315r con tiempos de exposición de 0.3 segundos. Las imágenes de difracción fueron procesadas y escaladas empleando el programa *iMOSFLM* y *SCALA* del paquete informático *CCP4*. La estructura fue resuelta por remplazo molecular con el programa *MOLREP* (Vagin & Teplyakov, 1997), utilizando como modelo la estructura de la porción Fc del IgG1 humano (código PDB: 1HZH). Una vez obtenidas las fases, se sometió a sucesivos ciclos de refinamiento con el programa *REFMAC* incluido en el conjunto de programas *CCP4* y la colocación de moléculas de agua se realizó con el programa *COOT* (Emsley & Cowtan,

2004). La validación de la estructura fue llevada a cabo con el programa Sfcchek (Vaguine et al., 1999) incluido en el conjunto de programas CCP4.

3.9.4 Ensayos de dinámica molecular del azúcar terminal

Los ensayos de dinámica molecular se llevaron a cabo con el fragmento Fc determinado por cristalografía (carente del ácido siálico terminal) y con una estructura homóloga de código PDB: 1MCO (sialilado α 2-3). Inicialmente, los campos de fuerza de la estructura de las proteínas fueron realizados con ff99SB (Hornak et al., 2006) y las cadenas glicánicas con Glycam_06 (Kirschner et al., 2008). Las estructuras preparadas fueron inmersas en cajas triclinicas con moléculas de agua TIP3P (Jorgensen et al., 1983). Los contornos fueron también añadidos manteniendo la electroneutralidad. Antes y después de añadir iones se realizaron dos minimizaciones de 10000 iteraciones cada una usando el método de pasos descendientes y gradiente conjugado. Todas las simulaciones de dinámica molecular fueron realizadas usando el paquete GROMACS 4.0.7 (Van Der Spoel et al., 2005). El tiempo de cada experimento fue de 2fs. Las condiciones de: número total de átomos, presión y temperatura se mantuvieron constantes. Las interacciones electrostáticas fueron evaluadas usando el método de malla-partícula de Ewald (*PME*) (Darden & Pedersen, 1993) con interacciones de Van der Waals cortadas a 14 Å. El sistema fue calentado y equilibrado de 50 K a 298 K. La longitud de los enlaces fue restringido usando el algoritmo de *LINCS* (Hess et al., 1997), y la geometría del agua fue restringida con *SETTLE* (Miyamoto & Kollman, 1992). La trayectoria fue evaluada en periodos de 1 ps.

RESULTADOS

4 RESULTADOS

4.1 DETERMINACIÓN ESTRUCTURAL DE LA ENDOLISINA Cpl-7

4.1.1 Resolución estructural del módulo catalítico de Cpl-7

4.1.1.1 Expresión y purificación de Cpl-7

La purificación de Cpl-7 recombinante inicialmente como proteína de fusión LSL_Cpl-7, se llevó a cabo a partir del lisado de 9 litros de cultivo; las fracciones obtenidas en el primer paso de purificación por cromatografía de afinidad, se observó una gran cantidad de la proteína de fusión LSL_Cpl-7 con un peso molecular correspondiente a la suma de ambas proteínas (≈ 50 kDa); a partir de aquí se unieron las fracciones enriquecidas con la proteína de fusión, donde se obtuvo un contenido total de proteína de 128.64 mg en 25 ml.

Una vez cortada la proteína de fusión con la proteasa TEV, se siguió con el segundo paso de purificación con una cromatografía de exclusión molecular, Superdex 75, en donde se obtuvo Cpl-7 a homogeneidad (figura 4.1), con un peso molecular correspondiente a 38 kDa. En el cromatograma se observan cuatro picos, el primero corresponde a una porción de la proteína de fusión que no se ha cortado; el segundo corresponde a Cpl-7, el siguiente corresponde a la proteasa y el último correspondiente a la lectina LSL. Se unieron las fracciones adecuadas de Cpl-7 pura de modo que se obtuvieron 30 mg totales, que se concentraron hasta $12 \text{ mg}\cdot\text{ml}^{-1}$ en tampón Tris-HCl 10 mM pH 8 y NaCl 100 mM, para los posteriores ensayos de cristalización.

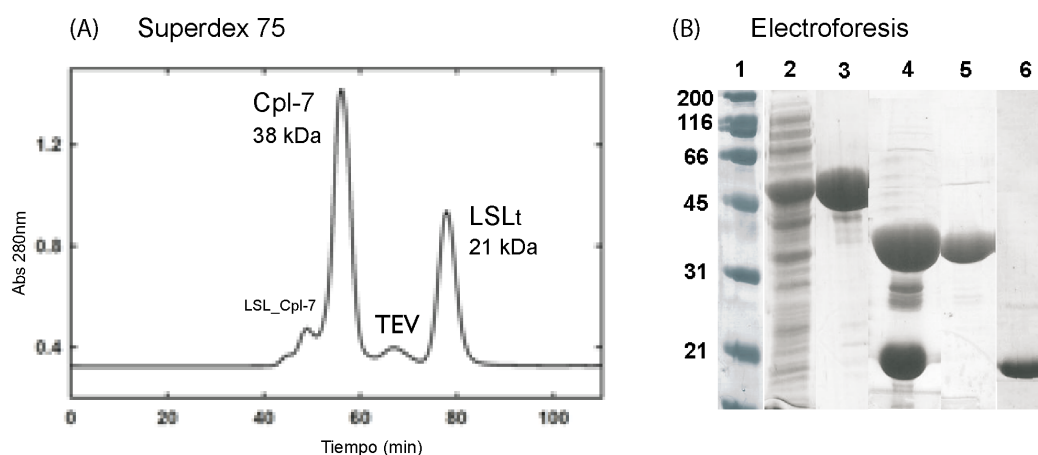


Figura 4.1. Purificación de Cpl-7. (A) Cromatograma Superdex 75. (B) Electroforesis en gel SDS-PAGE al 12%, correspondiente a las distintas etapas del proceso de purificación. El canal 2 corresponde al extracto celular, el canal 3 representa el primer paso de purificación con sepharosa 4B, el canal 4 muestra el resultado de la reacción de corte con la proteasa TEV y los canales 5 y 6 corresponden a la cromatografía de exclusión molecular Superdex 75 para el pico de Cpl-7 purificada y la LSL, respectivamente. La movilidad de los marcadores de peso molecular (Mw) se indican por orden creciente: miosina, β -galactosidasa, fosforilasa, albúmina bovina, ovoalbúmina, anhidrasa carbónica, el inhibidor de tripsina, lisozima y aprotinina.

En base a los resultados obtenidos no cabe duda de que el rendimiento de la purificación, es satisfactorio, obteniéndose una elevada cantidad de Cpl-7 pura. La homogeneidad de la muestra se verificó a través del espectro de masas por MALDI-TOF (figura 4.2), en donde observamos un único pico con una masa aparente de 38576 Da.

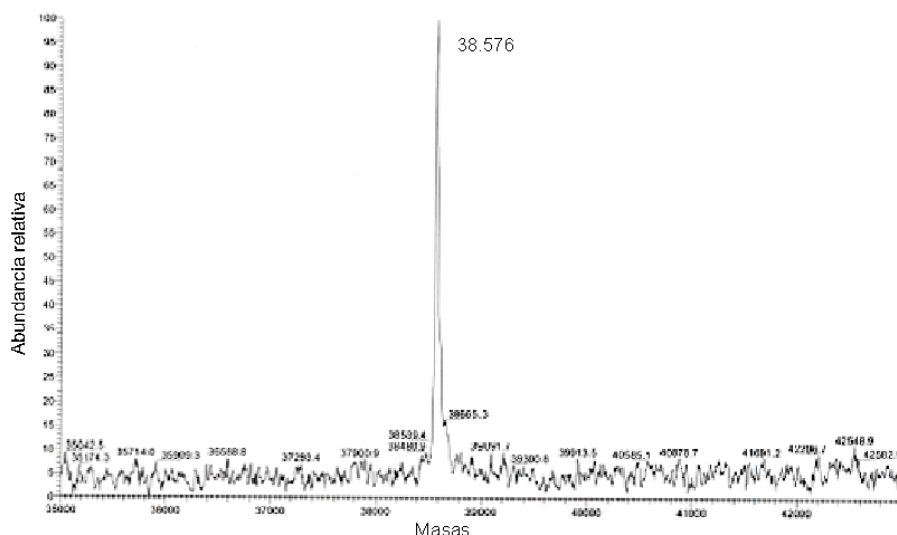


Figura 4.2. Perfil de espectrometría de masas (MALDI-TOF) de la lisozima Cpl-7 purificada. Los resultados muestran un pico mayoritario con una masa de 38576 kDa y otros minoritarios debidos a la formación de aductos con la matriz durante la ionización.

4.1.1.2 Cristalización de Cpl-7

Las condiciones de cristalización se establecieron bajo técnicas de alto rendimiento, ensayando con diferentes condiciones iniciales. Al cabo de 4 meses se observaron cristales de pequeño tamaño en dos condiciones, (i) 0.2 M de formiato de sodio, 20% PEG 3350 y (ii) 0.2 M de bromuro de sodio, 0.1 M de Bis-Tris propano pH 6.5 y 20% PEG 3350 (pertenecientes al PACT+Suite de QIAGEN, N° 54 y N° 62 respectivamente), sobre las microplacas de cristalización, con dimensiones 0.1 x 0.08 x 0.1 mm (Figura 4.3).

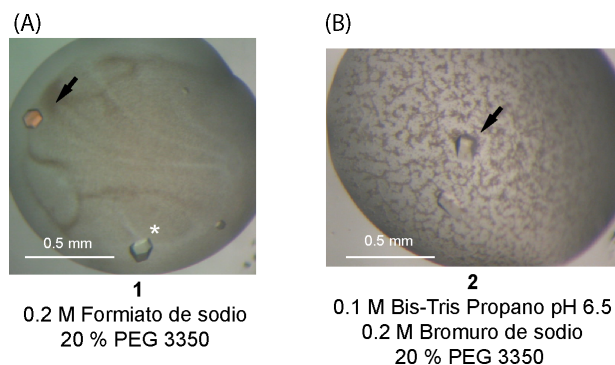
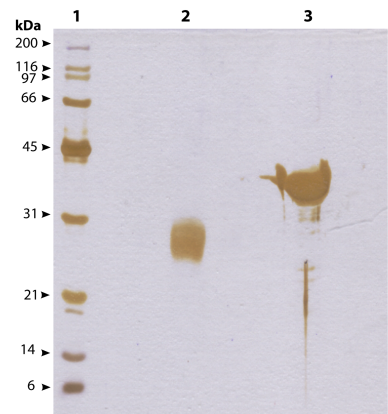


Figura 4.3. Cristalización de la endolisina Cpl-7. (A) Condición 1 de cristalización. (B) Condición 2 de cristalización. Las flechas indican los cristales utilizados en los ensayos de cristalización. (*) Indica el cristal utilizado en el ensayo de la integridad de la proteína en el cristal.

4.1.1.3 Integridad de la proteína Cpl-7 en los cristales

Debido a que los cristales crecieron al cabo de 4 meses decidimos realizar una prueba de la integridad de la proteína en los cristales obtenidos, sometiendo un cristal (indicado con * en la figura 4.3-A) a electroforesis en SDS-PAGE seguido de la tinción de plata (Switzer *et al.*, 1979), y comparándolo con una muestra control de Cpl-7 pura del mismo lote utilizado en los ensayos de cristalización. Al revelar el gel (figura 4.4), notamos claramente que la proteína dentro del cristal estaba parcialmente cortada, presentando un peso molecular de 28 kDa aproximadamente, el cual podría ser compatible con el módulo catalítico de Cpl-7.

Figura 4.4. Análisis de la integridad de Cpl-7 en el cristal por electroforesis en SDS-PAGE seguido de la tinción con plata. Canal 1, se indican los marcadores de peso molecular en orden decreciente en kDa a la izquierda (miosina, β -galactosidasa, fosforilasa, albúmina bovina, ovoalbúmina, anhidrasa carbónica, inhibidor de tripsina, lisozima y aprotinina.). Canal 2 cristal de Cpl-7 disuelto. Canal 3 muestra de Cpl-7 pura.



4.1.1.4 Medida de los datos de difracción de Cpl-7

A pesar de las pequeñas dimensiones de los cristales, se logró recoger un conjunto de datos a 2.4 Å de resolución, de los cristales de la condición 1 (crioprottegidos en glicerol 20%). Los patrones de difracción obtenidos fueron de excelente calidad (figura 4.5-A) y se recogieron en una fuente de rayos X de ánodo rotatorio, en el Departamento de Cristalografía y Biología Estructural del Instituto Química Física “Rocasolano” del CSIC. El resultado del indexado indicó que los cristales pertenecían al sistema cristalino tetragonal I422, presentando unas dimensiones de celdilla de $a = 127.93$ Å, $b = 127.93$ Å, $c = 83.03$ Å. Las estadísticas del procesado y escalado de los datos de difracción se listan en la tabla de la figura 4.5-B.

Teniendo en cuenta la masa molecular del módulo catalítico de la endolisina Cpl-7 y el volumen de la celdilla unidad, los datos de difracción sugieren la presencia de una molécula en la unidad asimétrica, con un contenido de solvente del 55.06% y un coeficiente de Matthews de $3.03 \text{ Å}^3 \cdot \text{Da}^{-1}$ (Matthews, 1968).

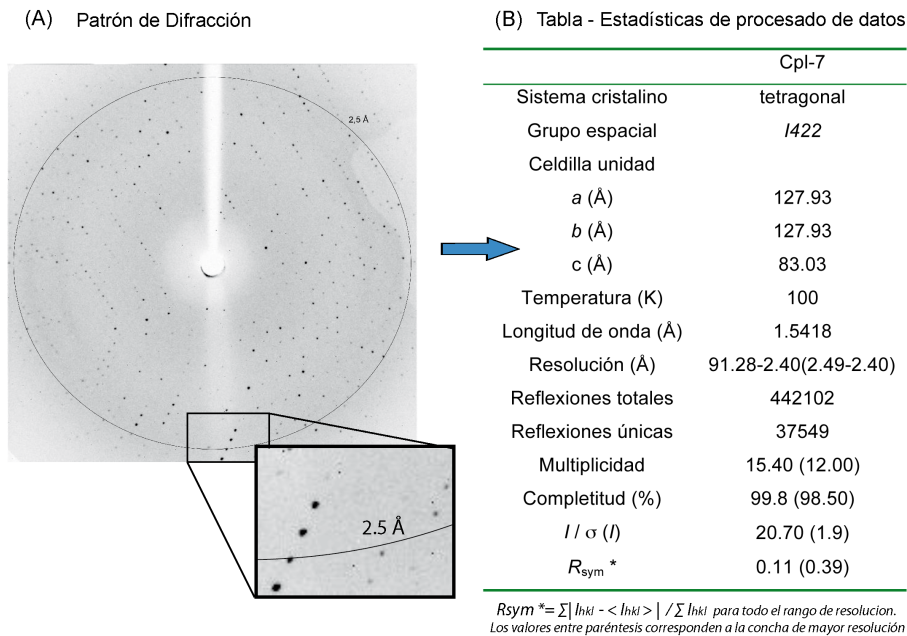


Figura 4.5. Medida de los datos de difracción del módulo catalítico de Cpl-7. (A). Patrón de difracción que muestra el limite de resolución a 2.5 Å (rango de oscilación 1°). (B). Tabla de las estadísticas del procesado y escalado de los datos de difracción.

4.1.1.5 Resolución estructural y refinamiento del módulo catalítico de Cpl-7

La resolución estructural se llevó a cabo mediante el Método de Reemplazo Molecular, utilizando como modelo inicial el módulo catalítico de Cpl-1 (Hermoso *et al.*, 2003), código PDB: 1H09. Este módulo presenta un 84% de homología de secuencia con el módulo catalítico de Cpl-7 (García *et al.*, 1990). El remplazo molecular fue realizado con el programa *MOLREP*, resultando una única solución contrastada para la función de rotación y translación. La correlación fue de 64.0% y el factor R de 38.0%. La inspección inicial del mapa de densidad electrónica reveló la presencia del módulo catalítico completo de Cpl-7 y la ausencia del módulo de unión a la pared celular. Obviamente la proteína nativa sufrió un proceso de degradación en la gota de cristalización dejando así que cristalizara el módulo catalítico (Silva-Martin *et al.*, 2010).

El modelo final del módulo catalítico de Cpl-7 fue refinado a 2.40 Å de resolución, lo que permitió modelar y ajustar perfectamente la cadena polipeptídica de 191 aminoácidos (residuos 2 - 192), dos moléculas de formiato y 207 moléculas de agua. Una vez refinada la estructura, se comprobó la estereoquímica a través del Diagrama de Ramachandran. El 89.4% de los residuos se encuentran dentro de las regiones favorables, el 10,6% de los residuos se observaron en regiones permitidas y ninguno en las zonas prohibidas. Las estadísticas finales del refinamiento y el diagrama de Ramachandran se muestran en la figura 4.6.

(A) Estadísticas del Refinamiento

| | Mod-Catalítico Cpl-7 |
|---|----------------------|
| Refinamiento | |
| Resolución (Å) | 91.28 - 2.40 |
| Nº Reflexiones | 13625 |
| Nº Reflexiones - R_{free} (% excluidas) | 1245 (5.3) |
| R (%) | 20.50 |
| R_{free} (%) | 25.20 |
| Modelo | |
| Moléculas / U. Asimétrica | 1 |
| Aminoácidos / U. Asimétrica | 191 |
| Nº Átomos | |
| Proteína | 1050 |
| Formiato | 6 |
| Agua (H ₂ O) | 207 |
| r.m.s.d. | |
| Longitud de enlace (Å) | 0.091 |
| Ángulo de enlace (°) | 1.557 |
| Valor promedio Factor B (Å ²) | 32.58 |

(B) Diagrama de Rachamandran

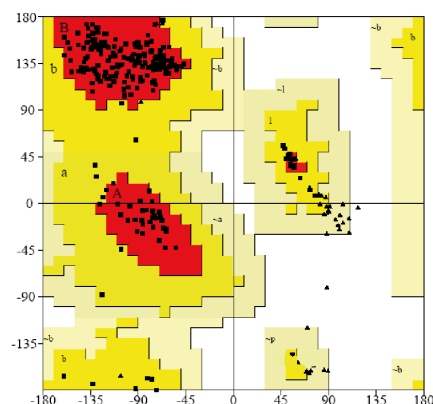


Figura 4.6. Refinamiento final de la estructura del módulo catalítico de Cpl-7. (A) Tabla de las estadísticas del refinamiento. (B) Validación del modelo a través del diagrama de Ramachandran.

4.1.1.6 Análisis de la estructura tridimensional del módulo catalítico de Cpl-7

El módulo catalítico de Cpl-7 presenta un plegamiento común con las glicosilhidrolasas de la familia 25 como el descrito para Cpl-1, y consiste en un barril irregular (α/β)₅ β ₃. En la figura 4.7 se muestran las 5 α -hélices características (en verde), las cuales bordean un cilindro central formado por 8 hojas β (en morado). El análisis estructural de la cavidad catalítica situada en la cavidad central (cavidad con superficie negativa), nos permite ver los residuos críticos en la actividad de la enzima, el Asp10 y Glu94. Estos residuos están 8.43 Å, distancia característica entre hidrolasas con mecanismo S_N1, con inversión de la configuración del carbono anomérico de la azúcar hidrolizado (Kuroki, et al

1999). Se deduce, por tanto, que Cpl-7 tiene un mecanismo catalítico del tipo S_N1 . Además podemos observar los residuos Asp92 y Asp182 que están formando puentes de hidrógeno de baja barrera (*low barrier hydrogen bonds*), con Glu94 (2.5 Å) y Asp10 (2.6 Å) respectivamente (figura 4.7), y que también están conservados con Cpl-1, lo que permitiría un libre tráfico de protones que promueve la regeneración del estado protonado de los residuos catalíticos principales (Olkhova *et al.*, 2007; Michel, *et al.*, 2001), maximizando así la velocidad de la reacción, hecho que se demostró para Cpl-1 (Hermoso *et al.*, 2003).

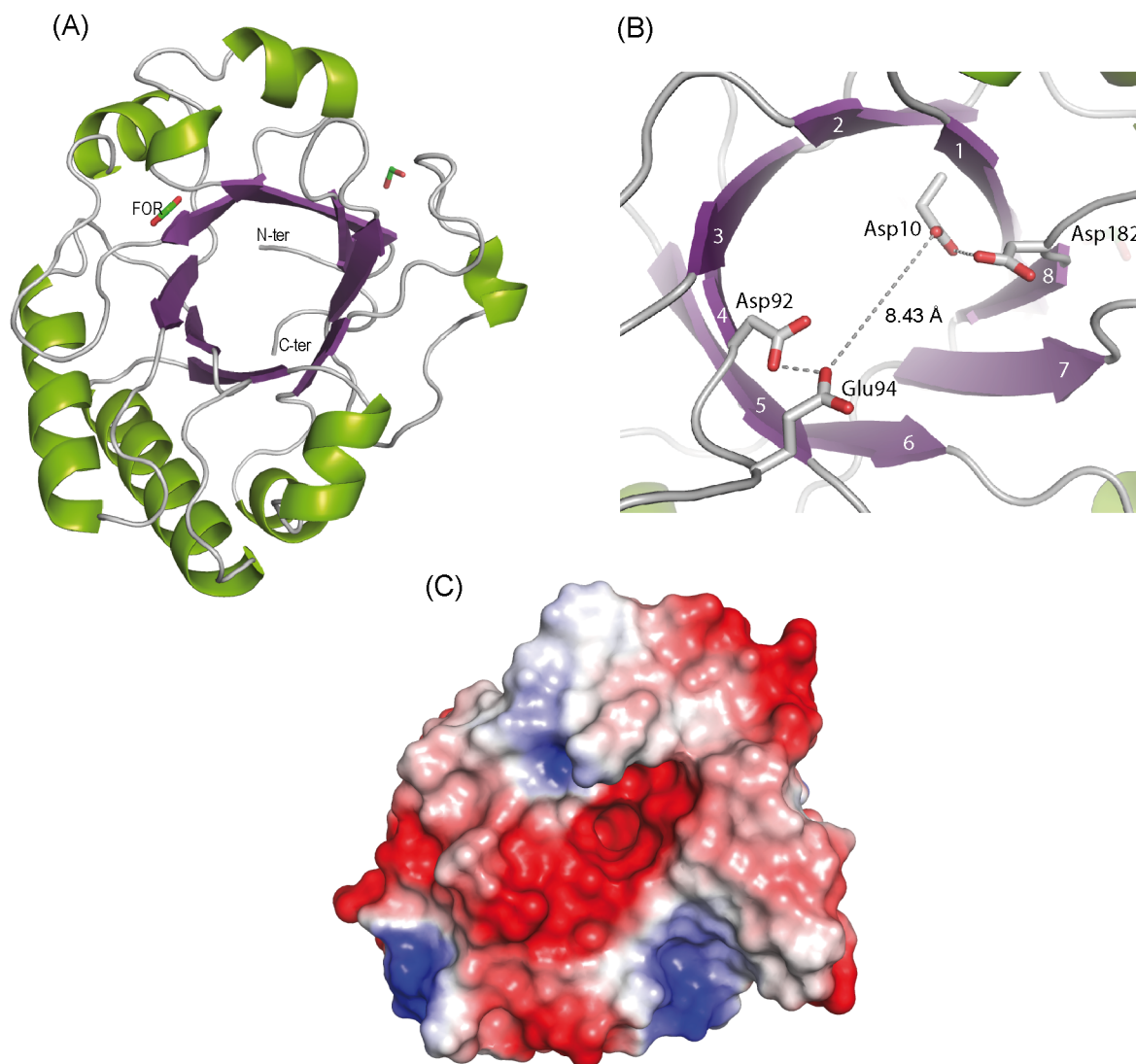


Figura 4.7. Estructura tridimensional del módulo catalítico de Cpl-7. (A) Estructura global del módulo catalítico mostrando un plegamiento característico $(\alpha/\beta)5\beta3$. (B) Cavity catalítica, se muestran los aminoácidos que estarían implicados en la hidrólisis, así como también se nota la hebra $\beta8$ antiparalela. (C). Superficie de potencial electrostático, en rojo cargas negativas, en azul cargas positivas y en blanco cargas neutras.

La superposición de la cadena principal de los módulos catalíticos de Cpl-7 y Cpl-1 arroja una desviación cuadrática media de 0.358 Å para 190 residuos (figura 4.8-A). Podemos estudiar las posibles interacciones entre el sustrato y Cpl-7 a través de lo ya descrito para Cpl-1. De este modo, en Cpl-7 podemos distinguir dos regiones de unión al peptidoglicano en función de las posiciones de la cadena glicánica, en la figura 4.9 se muestra la superposición con Cpl-1 en complejo con un fragmento de peptidoglicano compuesto por un disacacárido pentapéptido (GlcNAc-MurNAc-(L-Ala-

D-Glu-L-Lys-D-Ala-D-Ala), (2S5P)₃). La superposición de las estructuras revela las posiciones de las cadenas glicánicas desde -2 a +3. El sitio de unión al peptidoglicano designado como 1 se muestra en amarillo, el sitio 2 en azul y los residuos catalíticos en verde (figura 4.8-B).

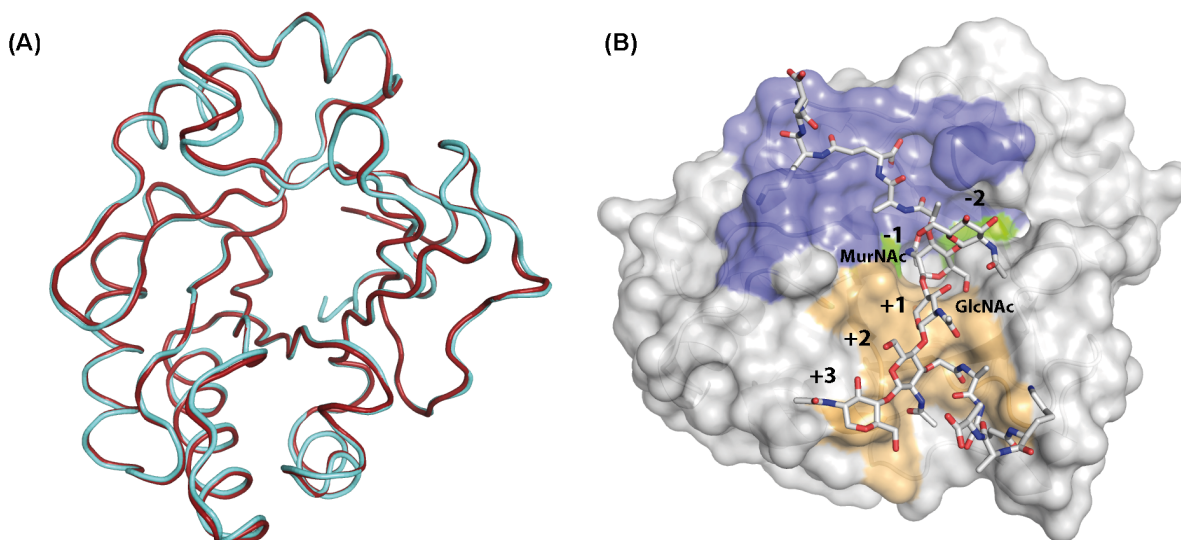


Figura 4.8. Superposición del módulo catalítico de Cpl-7 y Cpl-1. (A) superposición de la cadena principal de Cpl-7 (rojo) y Cpl-1 (azul) (B) Superficie de interacción de Cpl-7 con un fragmento de PG (2S5P)₃, por superposición con el modelo Cpl-1: (2S5P)₃, se señalan las posiciones de los azúcares de +3 a -2, las cadenas salientes son la parte peptídica del peptidoglicano.

Al comparar ambas estructuras notamos un cambio de conformación a nivel de la tirosina 127. En la estructura de Cpl-7 este residuo se encuentra en posición cerrada en comparación con el complejo Cpl-1:(2S5P)₂, bloqueando el acceso al sitio del sustrato (figura 4.9-A). Todos los aminoácidos que componen ambos sitios de unión al PG están conservados en Cpl-7, excepto dos aminoácidos (Asp96His y Arg63Tyr) en el sitio 2 de unión al PG. En lugar del puente salino formado por la Asp96 y Arg63 en Cpl-1, se forma una interacción de apilamiento entre el Tyr63 y la His96 en Cpl-7, a pesar de estas mutaciones la estructura no sufre grandes cambios en esta zona (figura 4.9-A).

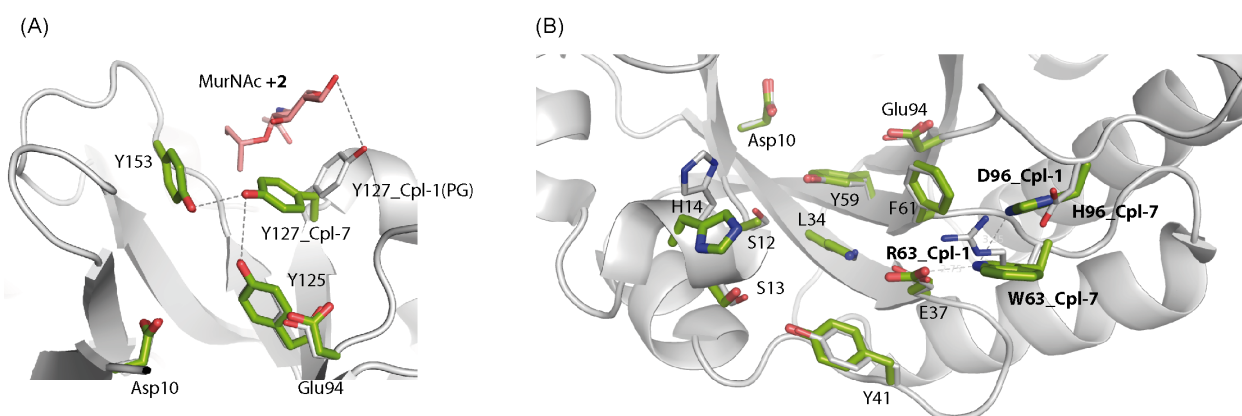


Figura 4.9. Superposición módulo catalítico de Cpl-7 y Cpl-1, mostrando en detalle las mutaciones importantes. (A) Sitio 1 de unión al PG, detalle del cambio de conformación de la Tyr127, el MurNAc en posición +2 se representa en rosa. (B) Sitio 2 de unión al PG. Residuos en verde corresponden a Cpl-7, residuos en blanco Cpl-1.

4.1.2 Determinación estructural del módulo de unión a la pared de Cpl-7

4.1.2.1 Expresión y purificación del extremo C-terminal de Cpl-7

Como ya se ha descrito anteriormente, la proteína Cpl-7 se divide en dos módulos estructurales, el módulo catalítico en el extremo N-terminal (188 aminoácidos) y el módulo de unión a la pared celular en el extremo C-terminal formado por 3 repeticiones idénticas, las dos primeras repeticiones de 48 aminoácidos cada una, y la tercera carece de los últimos seis aminoácidos. Estos dos módulos están conectados por una región de 16 aminoácidos (figura 4.10), sin estructura secundaria definida (según los programas de predicción). Debido a que el conector entre ambos módulos posiblemente sea bastante flexible, decidimos clonar el módulo de unión a la pared a partir de la Asn196, que se sitúa a la mitad del conector (figura 4.10), para evitar problemas de inestabilidad estructural que pudieran afectar los procesos de cristalización.

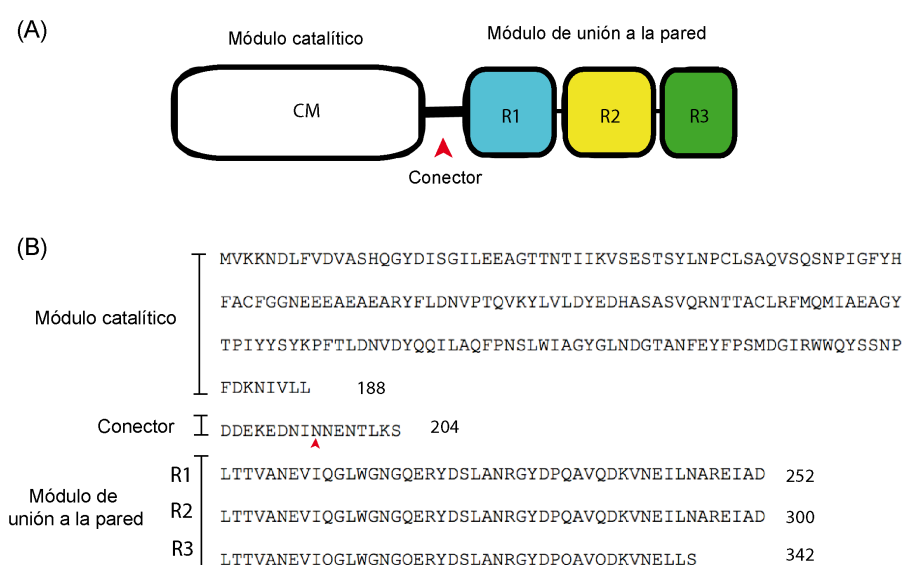


Figura 4.10. Representación de la disposición modular de Cpl-7. (A) El esquema muestra el módulo catalítico (blanco), módulo de unión a la pared diferenciado por repeticiones (azul, amarillo y verde) y el conector señalado con una punta de flecha roja. (B) Secuencia de aminoácidos en Cpl-7, se señala el Asp196 a partir de la cual se clonó el módulo de unión a la pared.

Tras la purificación del módulo de unión a la pared (CWBM) de Cpl-7, se logró concentrar la muestra a $20 \text{ mg} \cdot \text{ml}^{-1}$ y $40 \text{ mg} \cdot \text{ml}^{-1}$, para los posteriores ensayos de cristalización. En la figura 4.11 se muestran los resultados del proceso de purificación a través de una electroforesis en gel. Si bien en un principio el módulo de unión a la pared fue expresado como proteína de fusión con un peso molecular de 32 kDa, luego del corte con la proteasa TEV obtuvimos el CWBM de Cpl-7 con un peso molecular 16 kDa. La homogeneidad de la muestra se verificó a través del espectro de masas por MALDI-TOF, en donde observamos un único pico con una masa aparente de 16300 Da, lo que es compatible con el peso molecular calculado y observado por electroforesis del CWBM de Cpl-7 (figura 4.11).

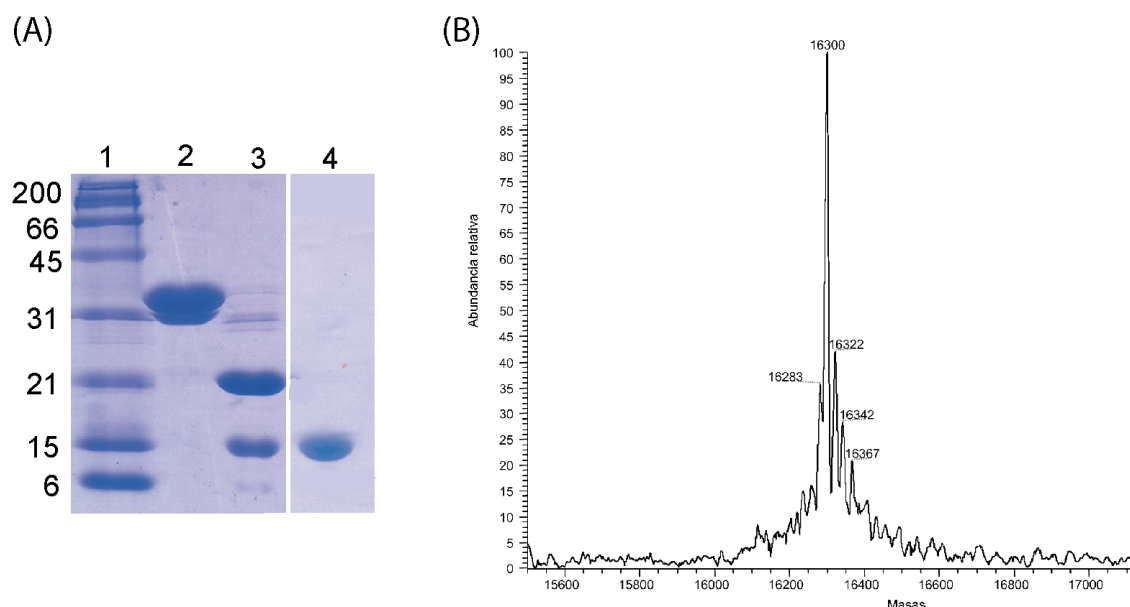


Figura 4.11. Purificación del CWBM de Cpl-7. (A) Electroforesis en SDS-PAGE (15%), donde se muestran las distintas etapas del proceso de purificación, así el canal 2 representa el primer paso de purificación con sefrosa 4B, la banda observada corresponde a la proteína de fusión, el canal 3 muestra el resultado de la reacción de corte con la proteasa TEV y el canal 4 corresponde a la proteína pura del CWBM de Cpl-7 purificada. La movilidad de los marcadores de peso molecular se indican por orden creciente a la izquierda. (B) Perfil de espectrometría de masas (*MALDI-TOF*) del dominio C-terminal de Cpl-7 purificada. Los resultados muestran un pico mayoritario con una masa de 161300 Da y otros minoritarios debidos a la formación de aductos con la matriz durante la ionización.

4.1.2.2 Cristalización del módulo de unión a la pared de Cpl-7

Tras los experimentos de cristalización iniciales (sección 3.7.2), se observó la formación de material cristalino, en la condición que contenía 2 M de sulfato de amonio y 0.1 M de Bis-Tris pH 5.5 y la condición que contenía 1.6 M de citrato tri-sódico, en estas dos condiciones se observaron cristales maclados y cristales de diminuto tamaño con alto grado de nucleación (figura 4.12-A).

(A) Condiciones Iniciales de cristalización

1: 0.1 M Bis-Tris pH 5,5
2 M Sulfato amónico

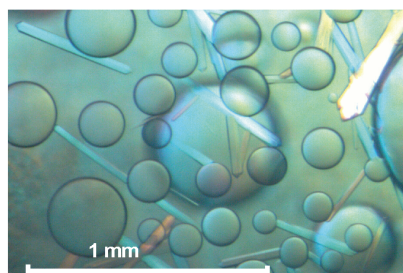


2: 1.6 M Citrato tri-sódico



gota: 250nl:250nl
reservorio: 70µl

(B) Escalado y Optimización



1.3 M Citrato tri-sódico

Aditivo: detergente
HEGA-8 - Octanoyl-N-hydroxyethylglucamide

gota: 2µl : 2.5µl : 0.5µl - reservorio - 500µl
proteína : reservorio : aditivo

Figura 4.12. Resultados de cristalización del módulo de unión a la pared de Cpl-7. (A) Condiciones iniciales de cristalización en donde se muestran los resultados obtenidos para la condición 1 (cristales maclados) y la condición 2 (cristales de pequeño tamaño y excesiva nucleación). (B) Se muestran los cristales obtenidos tras el escalado y la optimización a partir de la condición 2 y añadiendo detergente como aditivo. Las proporciones de gota se señalan para cada ensayo.

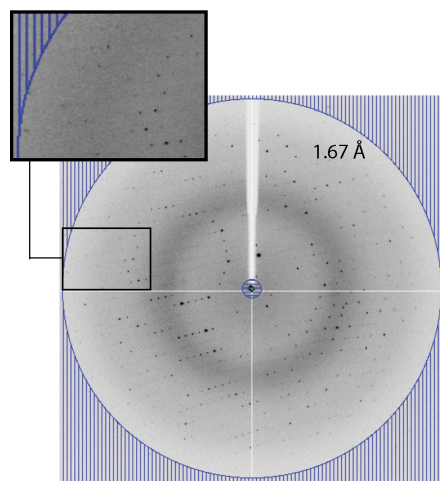
Se realizó un barrido de ambas condiciones de cristalización así como también variaciones en la concentración de la muestra proteica, consiguiendo con éxito el escalado de los cristales desde las gotas de 500 nl a 1 μ l, pero no se logró conseguir monocristales de buen tamaño y calidad. Por esta razón realizamos ensayos de cristalización agregando detergentes como aditivos a las gotas de cristalización.

Finalmente se obtuvieron cristales de excelente tamaño y calidad en un periodo de 12 días al usar el detergente HEGA-8 (Octanoyl-N-hydroxyethylglucamide). Los cristales presentaron una forma de prisma con unas dimensiones aproximadas de 0.7 x 0.1 x 0.1 mm, (figura 4.12-B). Estos cristales fueron utilizados para tomar datos de difracción, utilizando como solución crioprotectora 20% de glicerol en solución madre.

4.1.2.3 Medida de los datos de difracción

Los cristales optimizados del módulo de unión a la pared de Cpl-7 nos permitieron medir un conjunto de datos a 1.67 Å de resolución, en la línea ID14-1 del *ESRF*. Los espectros de difracción obtenido con estos cristales era de muy buena calidad (figura 4.13-A). Los cristales pertenecen al sistema cristalino hexagonal y presentan unas dimensiones de celdilla $a = b = 58.40$ Å, $c = 28.50$ Å, $\alpha = \beta = 90.0^\circ$, $\gamma = 120.0^\circ$. Los datos se procesaron con un grupo espacial $P6_1$. Los resultados del procesamiento se recogen en la figura 4.13-B.

(A) Patrón de Difracción



(B) Tabla - Estadísticas de procesamiento de datos

| C-terminal Cpl-7 | |
|----------------------|------------------------|
| Sistema cristalino | hexagonal |
| Grupo espacial | $P6_1$ |
| Celdilla unidad | |
| a (Å) | 50.40 |
| b (Å) | 50.40 |
| c (Å) | 28.55 |
| Temperatura (K) | 100 |
| Longitud de onda (Å) | 0.933400 |
| Resolución (Å) | 28.54-1.67 (1.76-1.67) |
| Reflexiones totales | 52758 (7259) |
| Reflexiones únicas | 4935 (701) |
| Multiplicidad | 10.70 (10.40) |
| Compleitud (%) | 99.60 (99.10) |
| $I / \sigma(I)$ | 30.90 (10.40) |
| R_{sym}^* | 0.05 (0.17) |

$R_{\text{sym}}^* = \sum |I_{hkl} - \langle I_{hkl} \rangle| / \sum I_{hkl}$ para todo el rango de resolución
Los valores entre paréntesis corresponden a la concha de mayor resolución

Figura 4.13. Medida de los datos de difracción del módulo C-terminal de Cpl-7. (A) Patrón de difracción que muestra el límite de resolución a 1.67 Å delimitado por el círculo azul, la ampliación de la imagen muestra con un recuadro en negro (rango de oscilación 1°). (B) Tabla de las estadísticas del procesamiento y escalado de los datos de difracción.

El volumen de la celdilla y su simetría eran incompatibles con la presencia del módulo de unión a la pared en su totalidad. Para comprobar la naturaleza del contenido de los cristales, se determinó la secuencia de aminoácidos que componen el extremo N-terminal, para este proceso se utilizaron cristales lavados y disueltos en el tampón de purificación. El resultado arrojó la secuencia peptídica: GEFNNENT. La secuencia obtenida está en acuerdo con el extremo N-terminal del módulo

de unión de Cpl-7 clonado. Los residuos GEF, corresponden a los tres aminoácidos que deja el corte con la proteasa TEV y la secuencia NNENT es el comienzo de la construcción en la posición 196 de Cpl-7. En conclusión la proteína en los cristales corresponde con el módulo de unión a la pared de Cpl-7.

La integridad de la proteína en los cristales se verificó a través del espectro de masas por MALDI-TOF, en donde observamos dos picos, uno de 12156 Da y otro de 6079 Da (figura 4.14), lo que indica que el cristal esta formado por dos especies de proteína. Como ya se mencionó anteriormente el módulo de unión a la pared de Cpl-7 esta constituido por 3 repeticiones. Si consideramos la masa molecular teórica de una sola repetición (≈ 6000 Da) y de dos repeticiones (≈ 12200 Da) y los resultados del espectro de masas, es muy posible que los cristales estén constituidos por una o dos repeticiones. En acuerdo con el volumen de la celdilla, para una repetición de 48 aminoácidos obtenemos un coeficiente de Matthews de $2.34 \text{ \AA}^3 \cdot \text{Da}^{-1}$ (Matthews, 1968) y un contenido de solvente del 47.40%, dicha celdilla no admite mas de una repetición, sólo es posible que los cristales contengan una sola repetición en la unidad asimétrica para el grupo espacial $P6_1$.

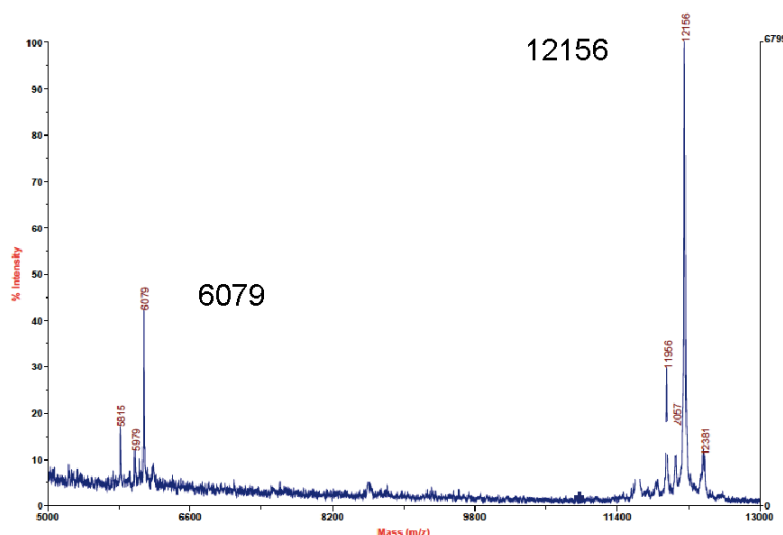


Figura 4.14. Perfil de espectrometría de masas (MALDI-TOF) del los cristales del módulo de unión a la pared de Cpl-7. Los resultados muestran dos picos mayoritarios con masas de 12156 Da y de 6079 Da y otros minoritarios debidos a la formación de aductos con la matriz durante la ionización.

4.1.2.4 Resolución estructural del módulo de unión a pared de Cpl-7

La determinación estructura del módulo de unión a la pared de Cpl-7 se llevó a cabo utilizando el programa *ARCIMBOLDO*, recientemente desarrollado (Rodríguez *et al.*, 2009). El éxito para determinar estructuras con el programa *ARCIMBOLDO* está sujeto a ciertos requerimientos: (1) la proteína debe tener un alto contenido en hélices alfa en su estructura secundaria, (2) la proteína a resolver no debe tener más de 100 aminoácidos y (3) los datos de difracción deben ser mejores de 2 Å de resolución. En este sentido el módulo de unión a la pared de Cpl-7 cumplía todas estas características. Así la mejor solución obtenida arrojó un modelo inicial de polialaninas, éste nos permitió calcular un mapa de densidad electrónica de alta calidad, con el cual se modeló y ajustó manualmente (programa *COOT*) los 48 aminoácidos de una repetición del módulo de unión a la pared

de Cpl-7. Una vez refinada la estructura, se comprobó la estereoquímica a través del Diagrama de Ramachandran. El 100% de los residuos se encontraron dentro de las regiones favorables. Las estadísticas finales del refinamiento y el diagrama de Ramachandran se muestran en la figura 4.15-A y 4.15-B.

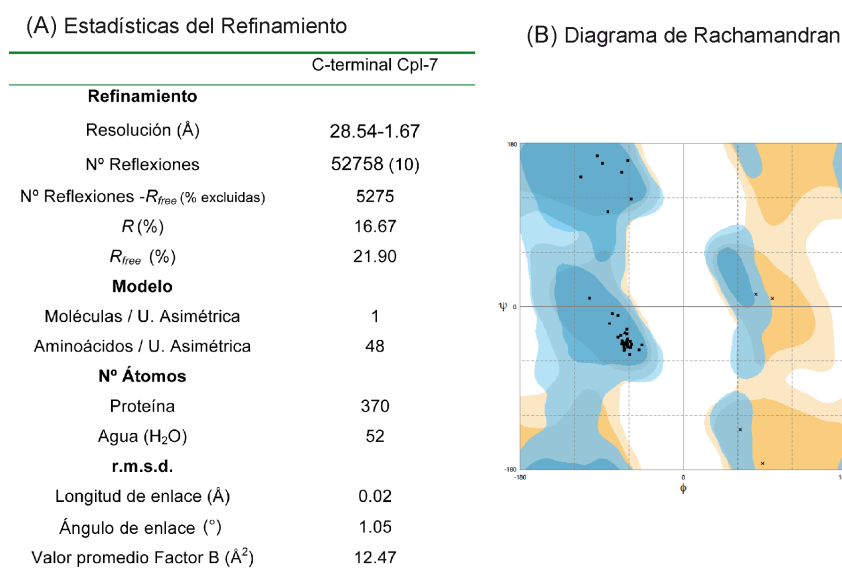


Figura 4.15. Refinamiento final de la estructura del módulo de unión a la pared de Cpl-7. (A) Tabla de las estadísticas del refinamiento. (B) Validación del modelo a través del diagrama de Ramachandran.

4.1.2.5 Análisis de la estructura tridimensional del módulo de unión a la pared de Cpl-7

La estructura tridimensional de la repetición del módulo de unión a la pared de Cpl-7, está formada por tres hélices- α ($\alpha 1$, $\alpha 2$ y $\alpha 3$), separadas por segmentos de 6 y 3 aminoácidos entre las hélices $\alpha 1$ y $\alpha 2$ respectivamente, además el bucle que conecta las tres repeticiones está formado por 6 aminoácidos, y lo definiremos como cL, (figura 4.16-A). Cabe destacar la abundancia de residuos ácidos en la superficie como puede observarse en el potencial electrostático de superficie, cargado negativamente (figura 4.16-B).

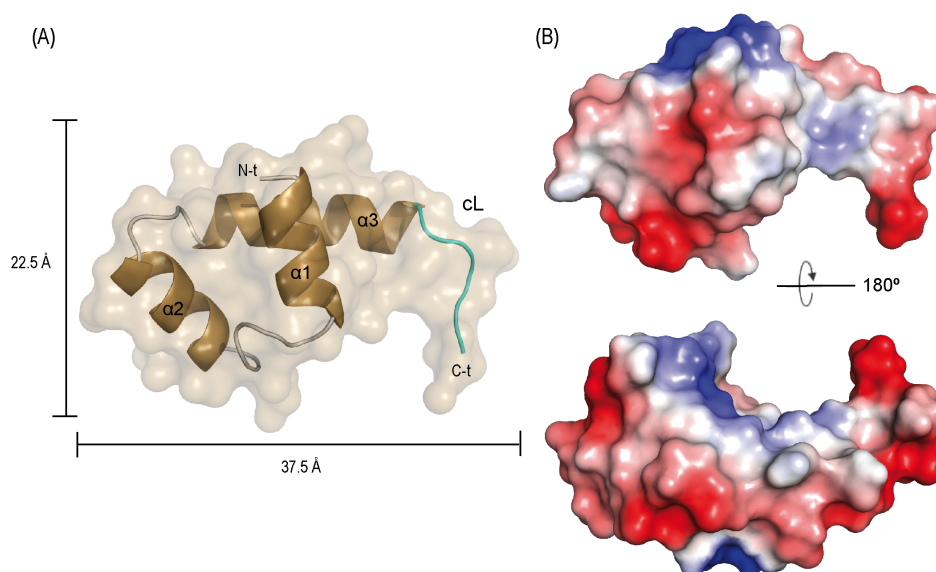


Figura 4.16. Estructura tridimensional de la repetición del módulo de unión a la pared de Cpl-7. (A) Se presenta una repetición la cual consta de tres hélices- α , $\alpha 1$, $\alpha 2$ y $\alpha 3$, (color dorado) y el fragmento cL que corresponde al conector entre las repeticiones (color turquesa). (B) Potencial electrostático de la superficie accesible al solvente (azul, positivo; rojo, negativo y blanco, neutro).

Esta estructura está fuertemente estabilizada por interacciones de naturaleza hidrofóbica entre las tres hélices- α (figura 4.17), a través de la Leu275 ubicada en el centro de la hélice α_2 , Val256 y Val260 (hélice α_2) y Val285 (hélice α_3), la Tyr280 ubicada en el bucle que conecta las hélices α_2 y α_3 funciona de tapa de este centro hidrofóbico, así como también el Trp265 ubicado en el bucle que conecta las hélices α_1 y α_2 . Adicionalmente la hélice α_1 se pliega sobre la hélice α_2 de forma anti-paralela, debido a la formación de un puente de hidrógeno fuerte (2.85 Å) entre el Glu258 de la hélice α_1 y el Trp265 del bucle que conecta ambas hélices- α (figura 4.17-A). Además podemos observar que las hélices α_1 y α_3 se estabilizan a través de un puente de hidrógeno (3.05 Å) formado por la Thr254 ubicada al comienzo de la hélice α_1 y la Lys288 ubicada en el centro de la hélice α_3 , la Lys288 a su vez forma un puente salino (2.55 Å) con el Glu291 (figura 4.17-A).

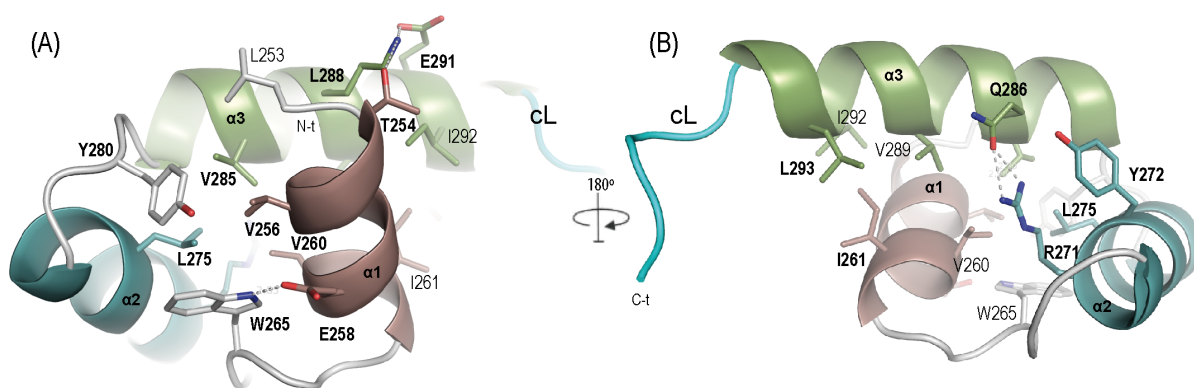


Figura 4.17. Análisis estructural de una repetición del módulo de unión a la pared de Cpl-7. (A) Muestra el centro hidrofóbico de la proteína que mantiene unidas a las tres hélices y los aminoácidos implicados. (B) Se representan los aminoácidos implicados en la estabilización de las hélices α_2 y α_3 . α_1 (marrón); α_2 (azul); α_3 (verde) y cL, bucle conector (turquesa).

Como puede verse en la figura 4.17-B, las hélices α_2 y α_3 están estabilizadas a través de un puente de hidrógeno fuerte (2.70 Å) formado por la Arg271 y la Gln286, y las interacciones hidrofóbicas, ya mencionas anteriormente, en el centro de las tres hélices. De este lado de la estructura también se observa una Tirosina (Tyr272) ubicada en el centro de la hélice α_2 que funciona de tapa del centro hidrofóbico al igual que la Tyr280 al otro lado de la estructura.

Tanto el módulo de unión a la pared de Cpl-7 como la repetición aislada, carecen de homólogos estructurales, sin embargo la estructura obtenida de esta repetición nos permitió el estudio comparativo con dos estructuras similares, RecX (código PDB:3C1D) y EBP (código PDB: 1H88). EBP es un regulador de la transcripción que tiene un módulo estructuralmente compatible con la estructura de la repetición del CWBM de Cpl-7 (según *Dali Server*). La superposición de ambas estructuras arrojó un r.m.s.d de 1.69 Å (figura 4.18-A). A pesar de que estas dos estructuras conservan el contenido de estructura secundaria y la disposición tridimensional de las hélices, no conservan las cargas en la superficie, mientras en Cpl-7 predominan las cargas negativas en EBP predominan aminoácidos cargados positivamente (figura 4.18-B), lo que le confiere a este factor de transcripción su capacidad para unirse al ADN.

RecX es una proteína modular ensamblada por repeticiones de tres hélices- α , mostrando una estructura alargada y ligeramente curva, su función está relacionada con la inhibición de la

transcripción a través de interacción directa con factores de transcripción. La superposición de la repetición del CWBM de Cpl-7 con las tres repeticiones por separado de RecX mostraron un r.m.s.d. de 4.04 Å, 1.51 Å y 4.90 Å, para las repeticiones 1, 2 y 3 en RecX (figura 4.18-C). Como puede observarse, únicamente la repetición central de RecX tiene un arreglo similar a la estructura tridimensional de la repetición de Cpl-7, aunque hay diferencias en la tercera hélice y en el conector. Como las repeticiones en RecX no son idénticas y en Cpl-7 si lo son, es poco probable que el arreglo entre las tres repeticiones en el CWBM de Cpl-7 sea similar a RecX. Además, la longitud de los conectores entre las repeticiones en RecX son más cortos (3 y 2 aminoácidos), en comparación con Cpl-7 (6 aminoácidos cada una), los conectores cortos en RecX restringen la flexibilidad y producen una disposición alargada de la repeticiones, mientras que en Cpl-7 los conectores largos serían más flexibles y por tanto el plegamiento del CWBM de Cpl-7 sería diferente a RecX (figura 4.18-C).

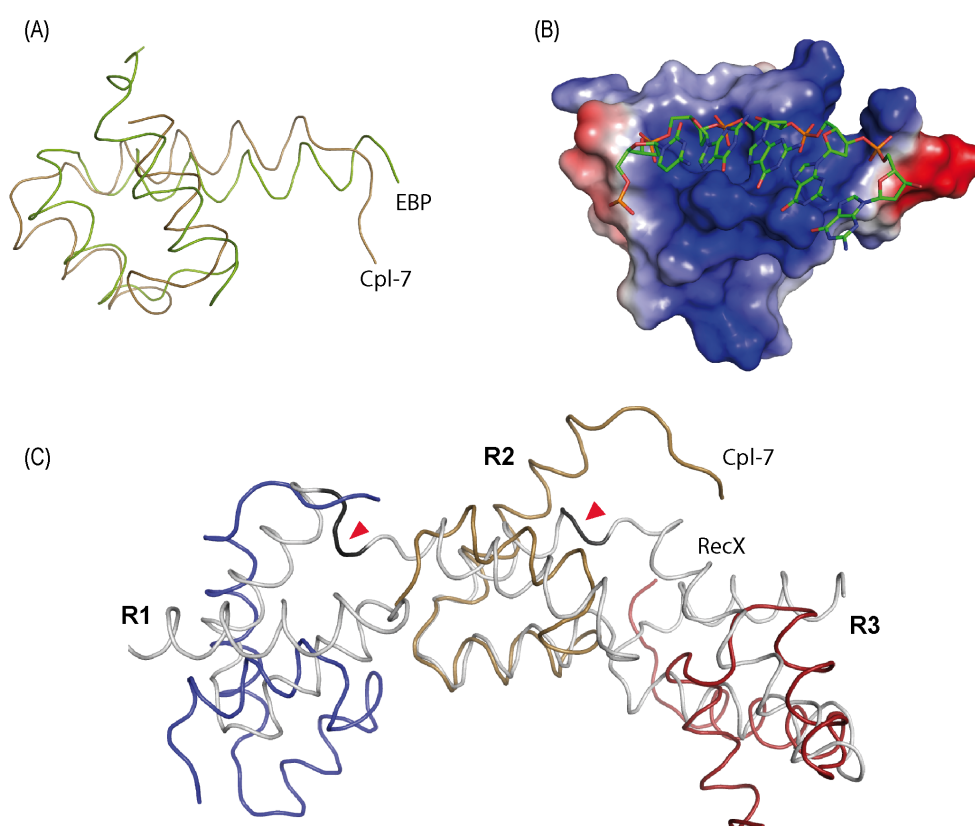


Figura 4.18. Comparación de estructuras relacionadas con el CWBM de Cpl-7. (A) Superposición de las cadenas polipeptídicas de la repetición de Cpl-7 (en marrón) con EBP (en verde). (B) Superficie de potencial electrostático de EBP en complejo con una hebra de ADN. (C) Superposición de las cadenas polipeptídicas de la repetición de Cpl-7 (en azul, marrón y rojo) con cada repetición de RecX (en gris), además se señalan los conectores entre las repeticiones de RecX con una punta de flecha roja (color negro en la estructura).

Al estudiar las moléculas relacionadas por simetría en el cristal, notamos que la densidad electrónica es continua, conectando el final de una cadena con el principio de la próxima. Como las repeticiones del CWBM de Cpl-7 son idénticas, este empaquetamiento cristalino nos permitió estudiar la disposición de tres repeticiones contiguas a lo largo de un eje de orden 3 en el cristal (figura 4.19).

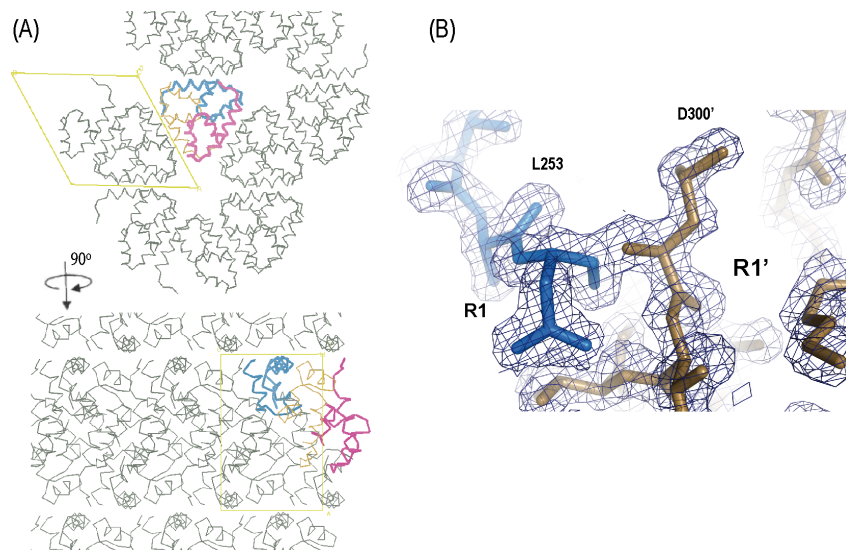


Figura 4.19. Empaquetamiento cristalino y densidad electrónica entre repeticiones. (A) Empaquetamiento de las moléculas en el cristal donde se muestran las cadenas contiguas en el eje de orden 3 en colores azul, amarillo y rosa. (B) Mapa de densidad electrónica (2Fo-Fc con un contorno de 1.5 σ) mostrando la conectividad entre L253 en el extremo N-terminal de la repetición (R1 en azul) y la D300' en el extremo C-terminal (R1' en marrón).

Las tres repeticiones contiguas (R1, R2 y R3), forman una estructura compacta, semicircular y helicoidal, ubicada a lo largo del eje de orden 3, con dimensiones aproximadas de 45.52 Å x 35.87 Å x 23.75 Å (figura 4.20-A). La estabilización entre las repeticiones vendría dada por una red de puentes de hidrógeno que se establecen entre las hélices $\alpha 2$ y $\alpha 3$ de la repetición R1 con las hélices $\alpha 1$ y $\alpha 2$ de la repetición R2. Es de gran importancia la formación de un puente salino (2.85 Å) entre la Arg248 en cL(1-2) el cual es parte de la R1 y el Glu258 ubicado al comienzo de la repetición R2 (hélice- $\alpha 1$), (figura 4.20-B). De la misma forma que se mantiene la estructura entre la R1 y la R2, se mantendrá entre la R2 y la R3, debido a que las repeticiones son idénticas. La hélice $\alpha 3$ de la R3 es más corta en comparación a las dos primeras repeticiones, esta última hélice no afectaría el empaquetamiento total de la estructura porque no estaría implicada en la estabilización.

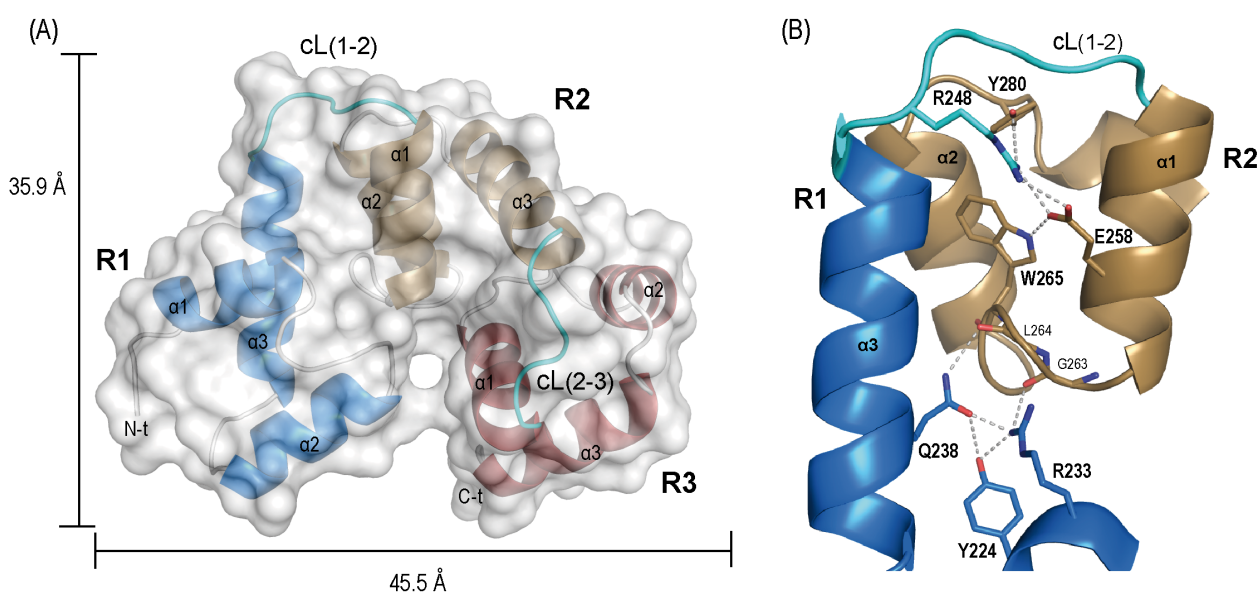


Figura 4.20. Posible estructura tridimensional del módulo de unión a la pared de Cpl-7. (A) Estructura tridimensional de tres repeticiones R1, R2 y R3, conectadas por los bucles cL(1-2) y cL(2-3) (B) Detalle de la estructura tridimensional donde se muestran las interacciones que estabilizarían las repeticiones entre sí, R1 (azul) y R2 (marrón), conectadas por los bucles cL(1-2) (celeste).

4.2 DETERMINACIÓN ESTRUCTURAL DEL RECEPTOR SIGN-R1

4.2.1 Cristalización de CRD_SIGN-R1

Las condiciones preliminares de cristalización se establecieron bajo técnicas de alto rendimiento. Al cabo de 7 días se observaron cristales de pequeño tamaño en diversas condiciones, estas condiciones tenían en común la naturaleza del agente precipitante (sales) a alta concentración (figura 4.21-A). Se escogieron dos de las condiciones iniciales para realizar el escalado de los cristales. Las condiciones seleccionadas fueron 0.1 M de Bis-Tris pH 5.5, 2 M de sulfato amónico y 1.6 M de Citrato tri-sódico (condición número 23 y 83 de JCSG+Suite QIAGEN respectivamente) (figura 4.21-A).

La optimización de las condiciones iniciales se llevó a cabo empleando el método de difusión de vapor en gota colgante en placas de 24 pocillos, variando las concentraciones del precipitante y el pH de la solución. Los mejores cristales se obtuvieron tras dos semanas y tenían forma de rombo (de dimensiones 0.06 x 0.06 x 0.01 mm) en 0.1 M Bis-Tris pH 5,5 y 1.7 M de sulfato amónico (figura 4.21).

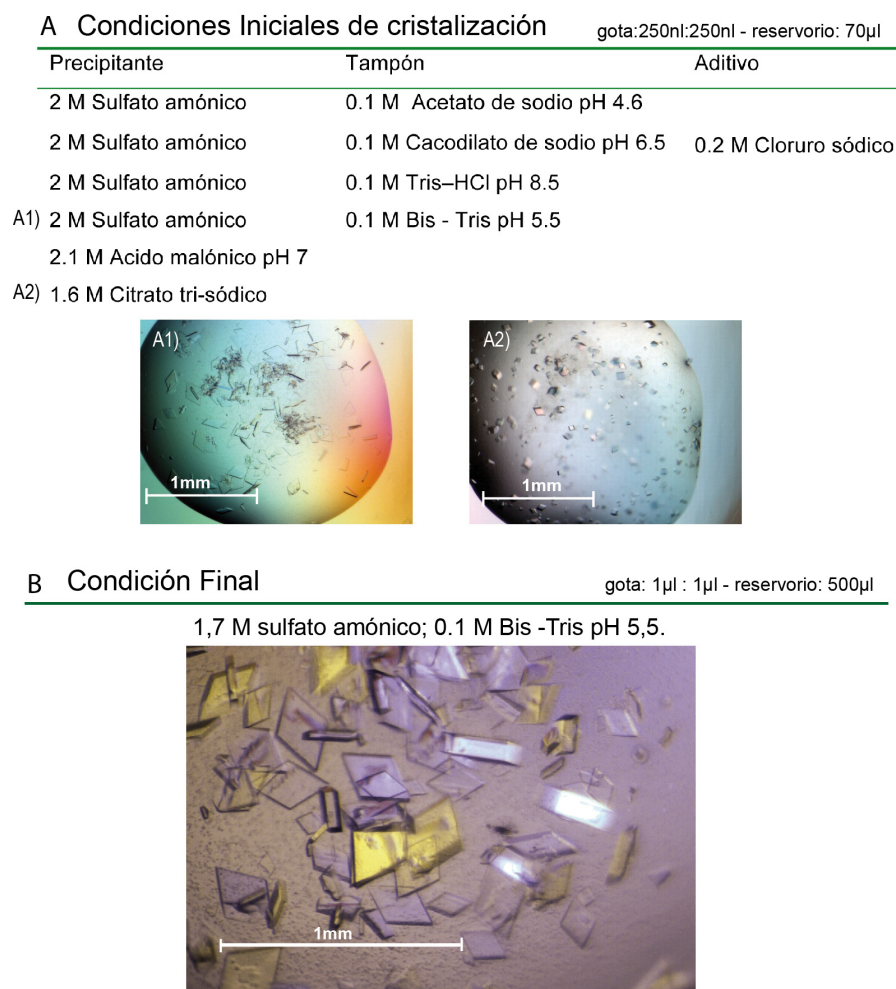


Figura 4.21. Cristalización de CRD_SIGN-R1. (A) Condiciones iniciales de cristalización, A1 y A2 condiciones y cristales obtenidos en los ensayos iniciales de cristalización en microplacas por el método de difusión de vapor en gota sentada. (B) Condición final, cristales escalados y optimizados a partir de la condición A1 por el método de difusión de vapor en gota colgante.

4.2.2 Medida de los datos de difracción

Los cristales fueron crioprottegidos en una solución que contenía una crio-sal (sulfato de litio saturado) mezclada con la solución de cristalización en una proporción 1:1 (figura 4.22-A). Se midió un conjunto de datos a 1.87 Å de resolución, (figura 4.22-B), en la línea ID23-1 del ESRF (Grenoble). El resultado del indexado indicó que los cristales pertenecían al sistema cristalino monoclinico, presentando unas dimensiones de celdilla de $a = 146.72$ Å, $b = 92.77$ Å, $c = 77.06$ Å y $\beta = 121.66^\circ$. Los datos se procesaron en el grupo espacial C2. Los resultados estadísticos del procesamiento y escalado de los datos de difracción se listan en la figura 4.22-C.

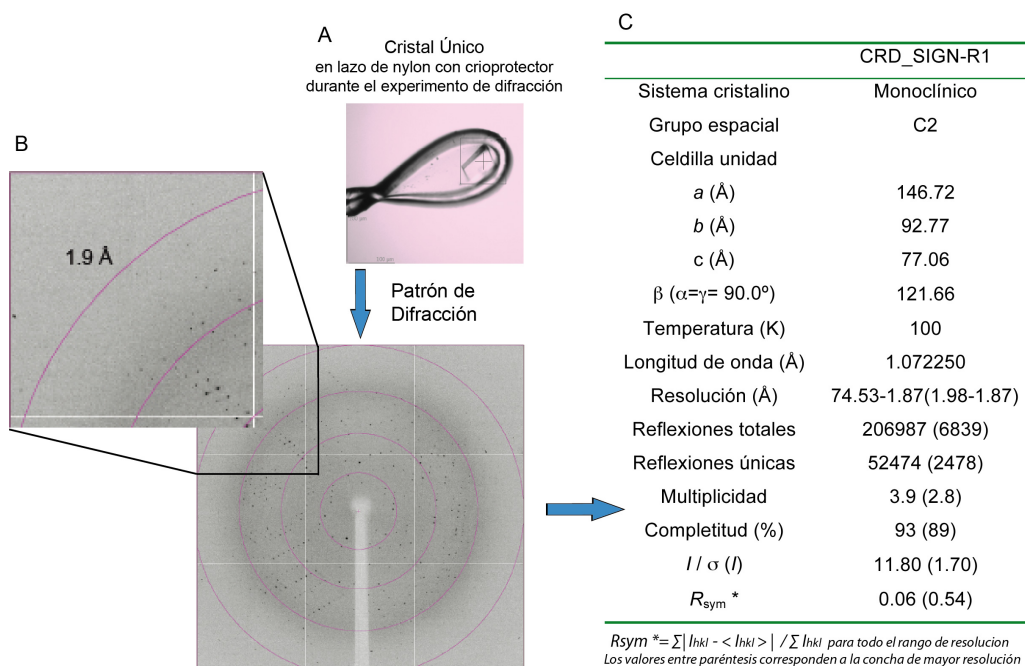


Figura 4.22. Medida de los datos de difracción. (A) Un cristal de CRD_SIGN-R1 montado en el lazo de nylon con solución crioprotectora. (B) Patrón de difracción. (C) Tabla de las estadísticas del procesado y escalado de los datos de difracción.

4.2.3 Resolución estructural y refinamiento

La resolución estructural se llevó a cabo mediante el Método de Reemplazo Molecular, utilizando como modelo estructural el dominio de unión a carbohidratos de DC-SIGN (Feinberg *et al.*, 2001), código PDB: IK91, que presenta un 68% de identidad de secuencia con el CRD_SIGN-R1. Teniendo en cuenta la masa molecular del CRD_SIGN-R1 y el volumen de la celdilla unidad, se obtiene un coeficiente de Matthews de $3.49 \text{ Å}^3 \cdot \text{Da}^{-1}$ (Matthews, 1968) y un contenido de solvente del 64.73 %, considerando cuatro moléculas en la unidad asimétrica. En el Reemplazo Molecular se obtuvieron cuatro soluciones contrastadas del resto, corroborándose así que efectivamente la unidad asimétrica contenía 4 cadenas del CRD_SIGN-R1. La correlación fue de 57.3% y el factor R de 39.7%.

El modelo final del CRD_SIGN-R1 fue refinado a 1.87 Å de resolución. El mapa de densidad electrónica nos permitió modelar y ajustar perfectamente cada cadena polipeptídica de 131 aminoácidos (residuos 161 - 292), un átomo de calcio por cada cadena polipeptídica, 6 moléculas de sulfato asociadas a cada cadena polipeptídica y 407 moléculas de agua. Una vez refinada la

estructura, se comprobó la estereoquímica a través del Diagrama de Ramachandran. El 95.18 % de los residuos se encuentran dentro de las regiones favorables, el 4.82 % de los residuos se observaron en regiones permitidas y ninguno en las zonas prohibidas. Las estadísticas finales del refinamiento y el diagrama de Ramachandran se muestran en la figura 4.23.

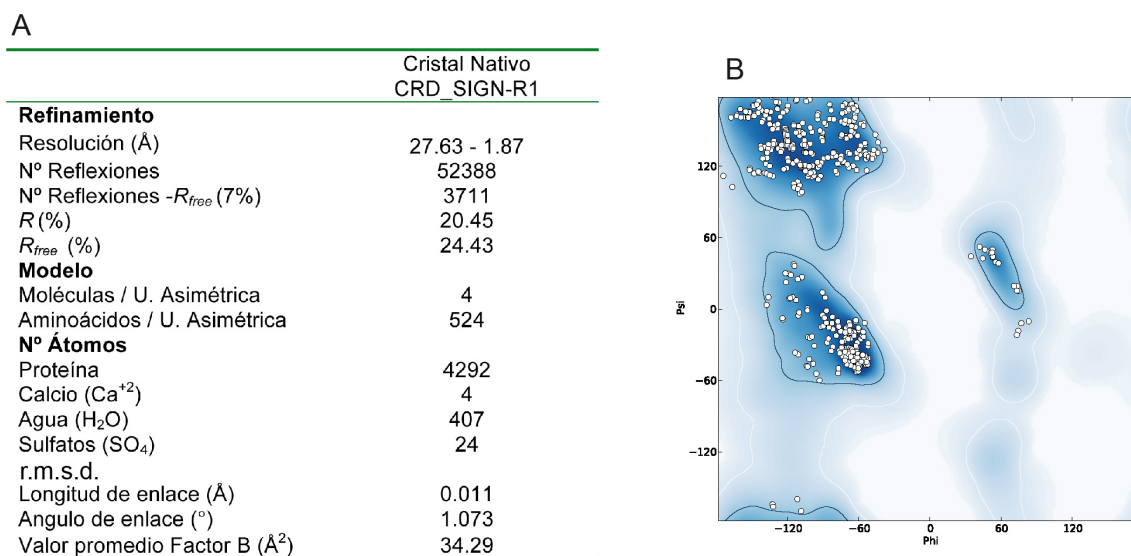


Figura 4.23. Refinamiento estructural de CRD_SIGN-R1. (A) Tabla de las estadísticas del refinamiento. (B) Validación del modelo a través del diagrama de Ramachandran.

4.2.4 Complejos cristalográficos con ligandos

4.2.4.1 Cristalización de los complejos

Para la obtención de los complejos cristalográficos del CRD_SIGN-R1 con carbohidratos, se ensayaron una amplia variedad de moléculas (tabla 2, apartado 3.11.3). De todos los complejos ensayados sólo se observó densidad electrónica para el sulfodextrano y el ácido siálico. Los complejos del CRD_SIGN-R1 : sulfodextrano y CRD_SIGN-R1 : Ácido siálico se obtuvieron por cocrystalización y por inmersión respectivamente. En los experimentos de cocrystalización el sulfodextrano se encontraba a una concentración final de 3% (v/v) y los cristales crecieron transcurridas tres semanas, con dimensiones similares a los cristales nativos. En los experimentos por inmersión, el ácido siálico se encontraba a una concentración final de 10 mM en la solución precipitante, los cristales nativos fueron sumergidos durante 12 horas.

Los ensayos de cristalización se repitieron a pH 8.5 (pH de mayor afinidad) y suplementados con 10 mM de calcio, para verificar la dependencia del pH o requerimiento del calcio, obteniéndose los mismos resultados.

4.2.4.2 Medida de los datos de difracción

Los cristales de ambos complejos fueron crioprotectados (sulfato de litio saturado : solución precipitante) y medidos en una fuente de radiación sincrotrón, obteniendo un conjunto de datos de difracción a 2.58 Å de resolución para el complejo con sulfodextrano y 2.80 Å de resolución para el complejo con ácido siálico (recogidos en la línea ID14-1 e ID23-2 respectivamente del ESRF). Ambos

cristales presentaron el mismo sistema cristalino y grupo espacial que los cristales nativos, con pequeñas variaciones en las dimensiones de la celdilla unidad. Los resultados obtenidos del procesamiento y escalado de los datos de difracción para los complejos se listan en la figura 4.24.

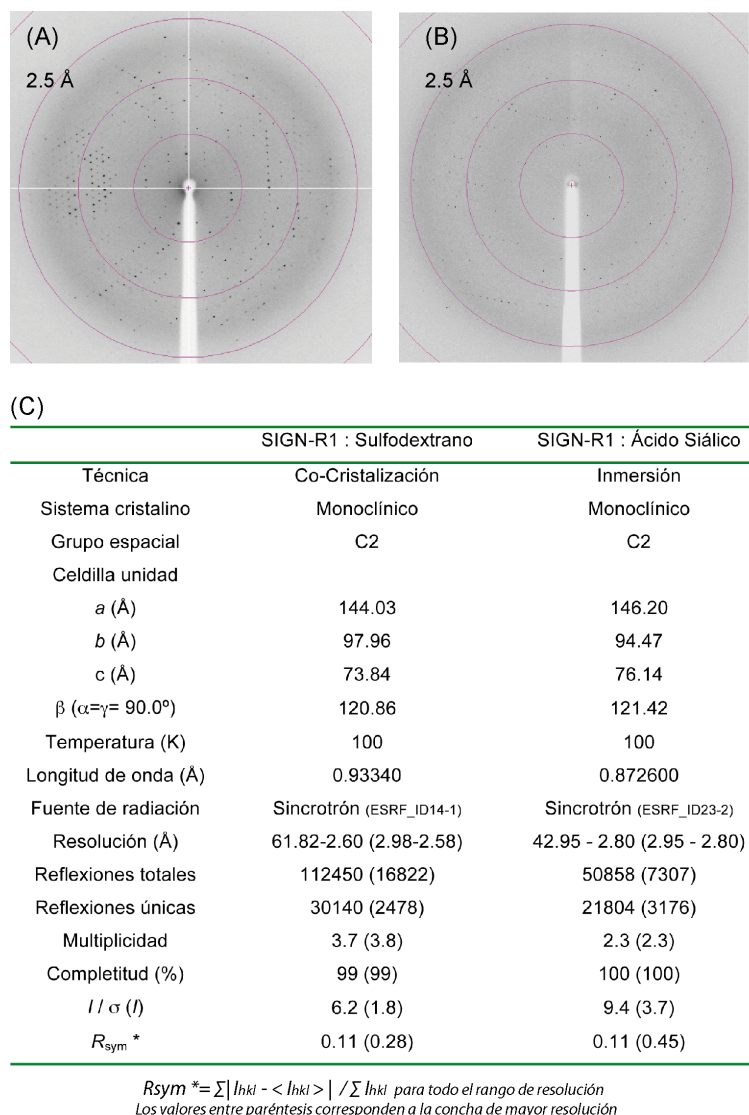


Figura 4.24. Medida de los datos de difracción. Patrones de difracción de los cristales de los complejos: (A) CRD_SIGN-R1 : Sulfodextrano y (B) CRD_SIGN-R1 : Ácido siálico. (C) Tabla de las estadísticas del procesado y escalado de los datos de difracción.

4.2.4.3 Resolución estructural y refinamiento

La estructura de los complejos se resolvió utilizando el Método de Reemplazo Molecular utilizando como modelo la estructura de SIGN-R1 nativa. En los dos casos se obtuvo la solución para las funciones de rotación y translación, con unos valores de correlación y factor R claramente diferenciadas del resto de las soluciones. La solución de reemplazo molecular para el complejo SIGN-R1 : sulfodextrano presentó una correlación de 75.4% y un factor R de 41.6% y la solución de reemplazo molecular para el complejo SIGN-R1 : ácido siálico arrojó una correlación de 67.9% y un factor R de 42.6%.

La estereoquímica de los modelos se verificó a través de los Diagramas de Ramachandran. Para el complejo con sulfodextrano el 93.50% de los residuos incluidos en la unidad asimétrica se encuentran dentro de las regiones favorables, el 6.50% en regiones permitidas y ninguno en zonas

prohibidas. Para el complejo con ácido siálico el 83.3% de los residuos se encuentran dentro de las regiones favorables, el 16.7% en regiones permitidas y ninguno en zonas prohibidas, los tres aminoácidos marcados con un punto rojo sobre el diagrama se encuentran sobre los límites pero no llegan a estar dentro de la zona prohibida. Las estadísticas finales del refinamiento de ambos modelos se muestran en la figura 4.25.

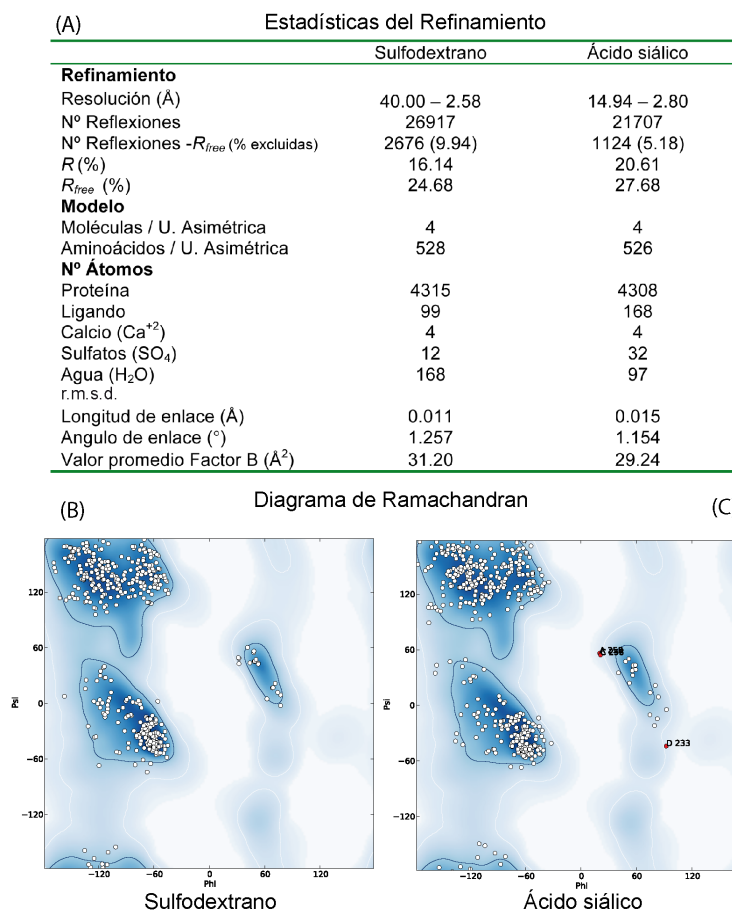


Figura 4.25. Refinamiento estructural de los complejos. (A) Tabla de las estadísticas del refinamiento. (B). Validación del modelo a través del diagrama de Ramachandran para el complejo CRD_SIGN-R1 : Sulfodextrano. (C) Validación del modelo a través del diagrama de Ramachandran para el complejo CRD_SIGN-R1 : Ácido Siálico.

4.2.5 Análisis estructural de SIGN-R1

4.2.5.1 Descripción General de la Estructura

La estructura del CRD_SIGN-R1 está formada por hélices α y láminas β . La hebra $\beta 2$ es la más larga en la estructura y divide la proteína en dos lóbulos característicos, lóbulo superior y lóbulo inferior (figura 4.26). El lóbulo superior está compuesto por una lámina β antiparalela, constituida por las hebras $\beta 2 - \beta 2' - \beta 3 - \beta 4$ y la hélice $\alpha 3_{10}$. En esta región se encuentra estabilizado el átomo de calcio principal (Ca1) (figura 4.26, esfera de color verde), dando lugar al sitio de unión de carbohidratos junto con la lámina β descrita anteriormente. Entre la hélice $\alpha 3_{10}$ y la hoja $\beta 3$ existe un bucle característico de este tipo de lectinas, llamado “Región del Lazo largo” LLR (de sus siglas en inglés: *long loop region*) y el bucle L2 que consiste en una horquilla entre las hebras $\beta 2$ y $\beta 2'$. El lóbulo inferior incluye los extremos N-terminal y C-terminal, estabilizados por tres puentes disulfuro (figura 4.26, representados como listones en color morado). Este lóbulo inferior está constituido por una

lámina β antiparalela formada por las hebras $\beta 0 - \beta 1 - \beta 1'$ y $\beta 5$, acompañados por las hélices $\alpha 1$ y $\alpha 2$. En la estructura observamos 6 moléculas de sulfato (S1-S6) que serán analizadas en detalle más adelante. Este plegamiento se conserva dentro de los miembros de la familia de las CTLD (de sus siglas en inglés: C-type lectin domain).

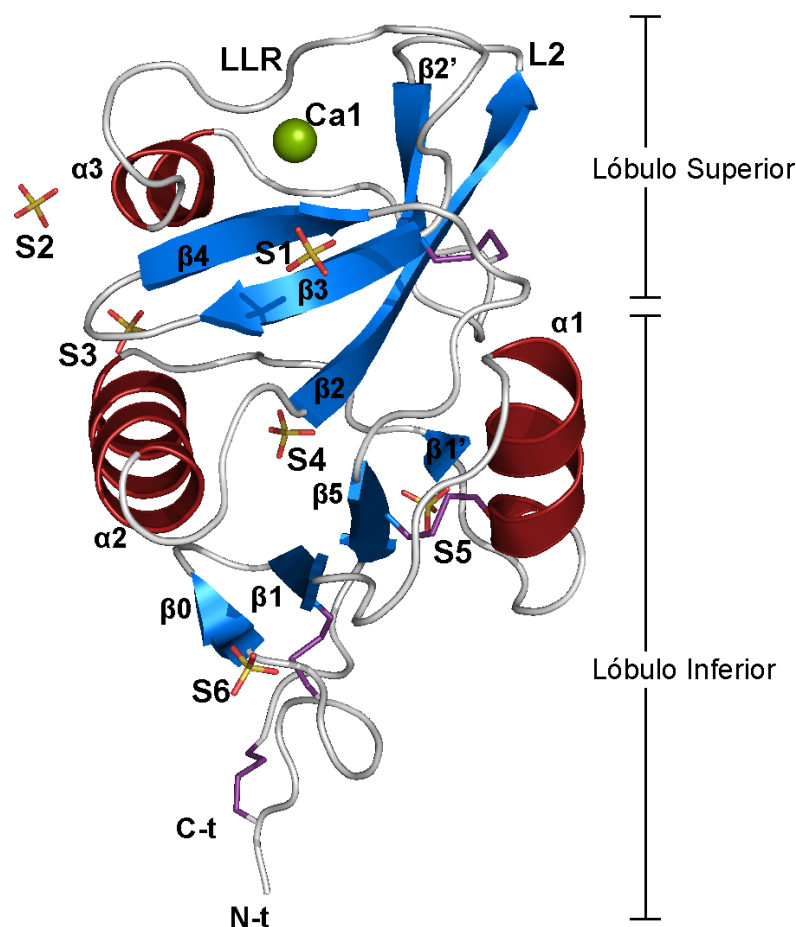


Figura 4.26. Estructura tridimensional del CRD_SIGN-R1. El átomo de calcio (Ca1) está representado como una esfera en verde. LLR, región de bucle larga. L2 bucle corto. Seis moléculas de sulfatos (S1-S6) están representadas como varillas, en color amarillo se representa el átomo de azufre y de color rojo los átomos de oxígeno. Los 4 puentes di-sulfuro están representados en color morado.

4.2.5.2 Sitio de unión a calcio

Los miembros pertenecientes a la familia de las lectinas dependientes de calcio, CTLD, requieren de la presencia de un átomo de calcio para el reconocimiento de los carbohidratos. Este átomo de calcio es designado como el principal (Ca1) y está presente en todas las estructuras resueltas, en algunos casos se observa la presencia de otros átomos de calcio designados como auxiliares o estructurales (Ca2 y Ca3) (Zelensky & Gready, 2003).

En la estructura cristalográfica del CRD_SIGN-R1, el calcio principal está estabilizado con una coordinación octaédrica, por los grupos carboxilato y carboxiamida de las cadenas laterales de Glu255, Asn257, Glu262, Asp274, el oxígeno carbonilo de la cadena principal de Asp274 y dos moléculas de agua (figura 4.27). Estos aminoácidos están conservados en todos los miembros de la familia. Los calcio auxiliares Ca2 y Ca3 están ausentes en SIGN-R1, a pesar de que las condiciones de cristalización estaban suplementadas con cloruro de calcio 10 mM.

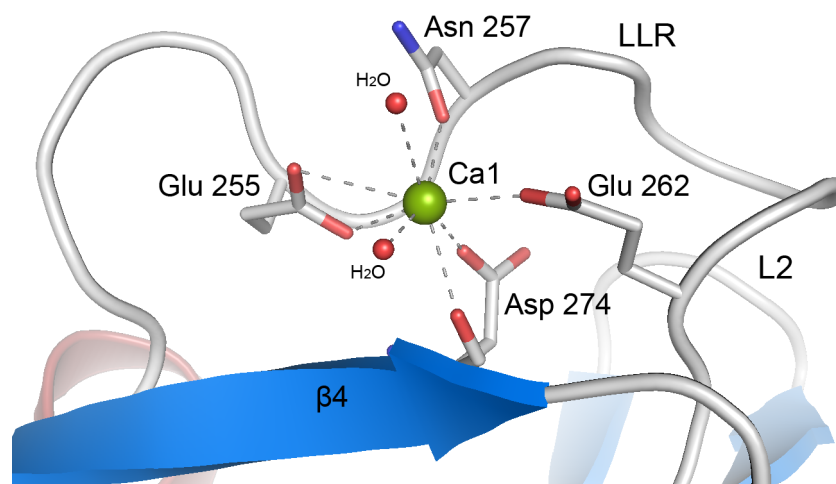


Figura 4.27. Estabilización octaédrica del átomo de calcio primario Ca1 en SIGN-R1. El átomo de calcio está representado con una esfera de color verde. La coordinación octaédrica está formada por las cadenas laterales de Glu255 (2.99 Å y 2.58 Å), Asn257 (2.46 Å), Glu262 (2.31 Å), Asp274 (2.36 Å), el oxígeno carbonilo de la cadena principal de Asp274 (2.59 Å), y dos moléculas de agua (2.16 Å, 2.82 Å), representadas con esferas de color rojo.

4.2.5.3 Comparación con otras lectinas homólogas de la familia – Ausencia de los calcios auxiliares

La comparación de las estructuras cristalográficas del CRD de SIGN-R1 con homólogos pertenecientes a la misma familia, como DC-SIGN (PDB: 1SL4) y DC-SIGNR (PDB: 1K9J), la forma apo de DC-SIGNR (PDB: 1XPH), (figura 4.28), indica dos diferencias bastante marcadas: la ya mencionada ausencia de los calcios Ca2 y Ca3 (característica que comparte con la forma apo de DC-SIGNR), y un cambio de la conformación del bucle LLR. En SIGN-R1, parte del bucle LLR está apuntando en dirección opuesta en comparación con el resto de las estructuras (señalado con una punta de flecha, figura 4.28, estructura de color azul), dejando al bucle LLR y al bucle L2 más cercanos entre sí que en el resto de las estructuras. Es importante destacar que a pesar de la ausencia de los calcios auxiliares, en SIGN-R1 el bucle L2 mantiene una conformación “cerrada”. En las estructuras de DC-SIGN y DC-SIGNR (Feinberg *et al.*, 2001; Guo *et al.*, 2004), la conformación cerrada del bucle L2 es estabilizada precisamente por los calcios auxiliares (estructuras roja y morado claro, figura 4.28). En la forma apo de DC-SIGNR, la ausencia de los calcios auxiliares conlleva a una flexibilidad tanto del bucle LLR como del bucle L2 provocando una conformación “abierta” (Snyder *et al.*, 2005; Chatwell *et al.*, 2008) (figura 4.28, estructura de color rosa).

Las diferencias observadas para el CRD_SIGN-R1, se pueden resumir como, cambio de la conformación del bucle LLR y estabilidad de la conformación “cerrada” del bucle L2 a pesar de la ausencia de los calcios auxiliares o estructurales supuestamente implicados en estabilizar la estructura de los homólogos. La comparación de secuencias de aminoácidos, reveló que los residuos que estabilizan los tres calcios están conservados (resaltados en gris figura 4.29). Pero también reveló que existen dos mutaciones relevantes en el L2 y LLR (Ala233 y Ile259) para SIGN-R1 (marcadas en color rojo figura 4.29).

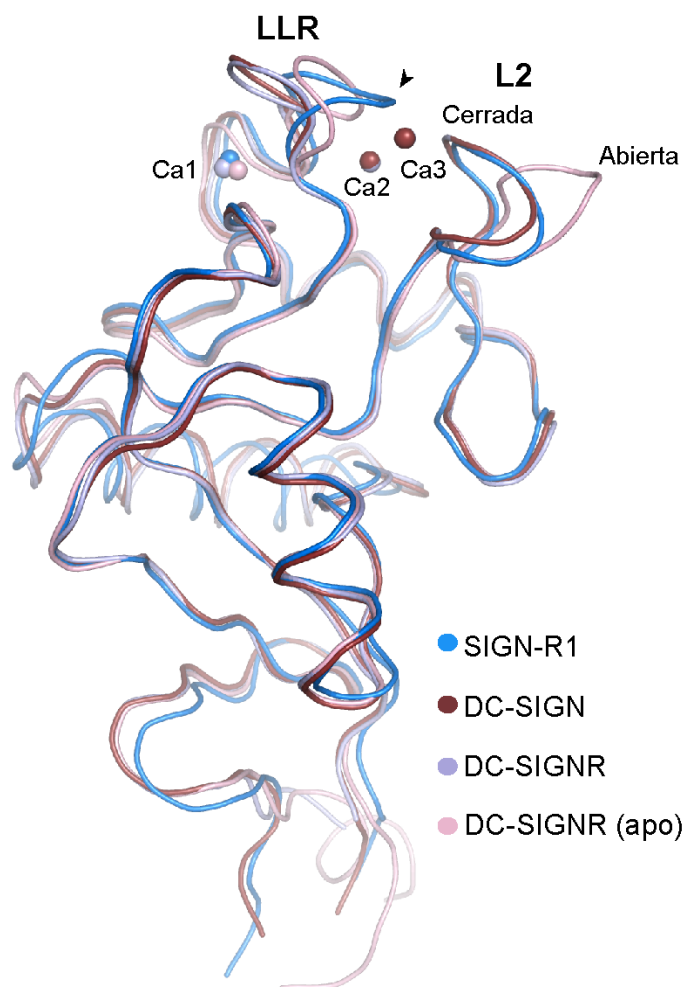
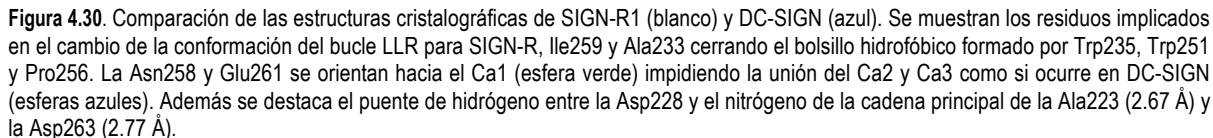
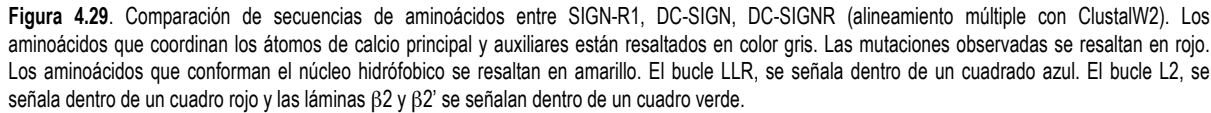


Figura 4.28. Superposición de la cadena principal de SIGN-R1 con sus homólogos. Se destaca la zona de unión a los calcio ubicada en lóbulo superior de la proteína. Los calcio primario (Ca1) y auxiliares (Ca2 y Ca3) se representan como esferas. CRD_SIGN-R1, presenta solo el calcio primario (color azul), se destaca el cambio de conformación de bucle LLR con una punta de flecha, quedando enfrentado al bucle L2 de conformación cerrada. DC-SIGN (PDB: 1SL4), presenta los 3 calcio (color rojo). DC-SIGNR (PDB: 1K9J), presenta los 3 calcio (color morado claro). La forma apo de DC-SIGNR (PDB: 1XPH), presenta el calcio primario (Ca1) mostrando una conformación abierta (color rosa).

En bucle LLR hay una mutación puntual en la posición 259 (numeración SIGN-R1), mientras que en SIGN-R1 es una Ile en DC-SIGN y en DC-SIGNR aparece una Val y una Ser, respectivamente (figura 4.29, en rojo). Yun Guo y colaboradores (2004) demostraron a través de experimentos de mutagénesis dirigida, que la Val351 y la Ser342 están involucradas en la especificidad del reconocimiento para los receptores DC-SIGN y DC-SIGNR, respectivamente. La presencia de Valina inclinaría la especificidad por carbohidratos ricos en fucosa, mientras que la presencia de Serina inclina la especificidad hacia carbohidratos ricos en manosa. En SIGN-R1 aparece en esta posición la Ile259, que a diferencia de DC-SIGN y DC-SIGNR no está expuesta sino que se orienta hacia un bolsillo hidrofóbico de la proteína, formado por Trp235, Trp251 y Pro256 (figura 4.29, residuos resaltados en amarillo). Esta distinta posición de la Ile259 explicaría la distinta orientación del bucle LLR para SIGN-R1 (figura 4.30), y consecuentemente SIGN-R1 presentaría una la distinta especificidad.



un puente de hidrógeno con el Asp263 (2.77 Å). Se han comprobado las interacciones con las moléculas relacionadas por simetría, y no influyen en la conformación de la estructura de SIGN-R1, lo que quiere decir que la conformación observada no es producto de un artefacto cristalográfico.

En resumen, a pesar de la absoluta conservación de los residuos implicados en la interacción con los tres calcio sólo el Ca1 está presente en SIGN-R1. La presencia de dos residuos hidrofóbicos en los bucles LLR y bucle L2, únicos en SIGN-R1, provocan cambios de conformación, en donde no es necesario la presencia de calcio auxiliares para mantener la estructura general. Estos cambios exponen residuos que originan un sitio de unión a carbohidratos completamente diferente del resto de las lectinas de su familia y que serían claves en la especificidad de SIGN-R1.

4.2.5.4 Presencia de moléculas de sulfato

Sorprendentemente seis moléculas de sulfato fueron observadas en la estructura cristalográfica del CRD de SIGN-R1 (S1-S6) (figura 4.31). El primero de ellos S1, se encuentra en la zona de reconocimiento a carbohidratos. El S1 está estabilizado por las cadenas laterales de Lys281, Ser275, Glu266 y dos moléculas de agua (figura 4.31). Además se observa un puente salino entre Lys281 y Glu266 (2.73 Å) que ayuda a estabilizar el S1.

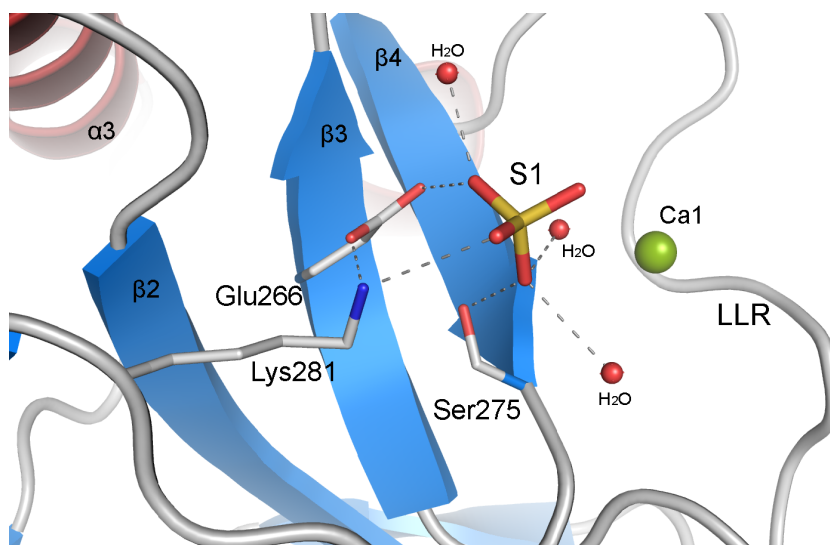


Figura 4.31. Estabilización de la molécula de sulfato 1 (S1) en SIGN-R1. El átomo de calcio está representado con una esfera de color verde. La molécula S1 está estabilizada por grupos carboxilato y amida de las cadenas laterales de Lys281 (3.35 Å), Ser275 (3.19 Å), Glu266 (2.92 Å) y dos moléculas de agua y dos moléculas de agua (2.54 Å, 2.68 Å y 2.84 Å), representadas como esferas de color rojo.

Las cuatro moléculas de sulfato siguientes (S2 - S5) están situadas en la cara opuesta a la zona de unión a carbohidratos (figura 4.26). Estas cuatro moléculas de sulfato sorprendentemente se disponen aproximadamente equidistante entre sí, espaciadas por 8.70 Å, 9.94 Å y 8.46 Å (figura 4.32).

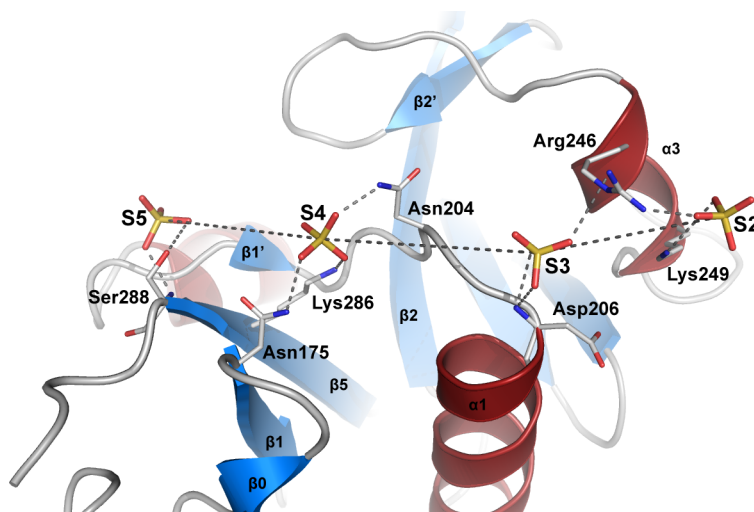


Figura 4.32. Disposición de los sulfatos S2-S5. La molécula S5 está interactuando con Ser288 (2.71 Å), el S4 está estabilizado por Asn175 (2.94 Å), Asn204 (2.74 Å) y Lys286 (3.19 Å). La molécula S3 está estabilizada por Asp206 (3.02 Å) y Arg246 (3.52 Å) y el S2 está interactuando con Arg246 (2.46 Å) y Lys249 (3.87 Å).

4.2.6 Descripción del sitio de unión a carbohidratos en SIGN-R1

4.2.6.1 Complejo CRD_SIGN-R1 : Sulfodextrano

El sulfodextrano es un polímero heterogéneo ramificado de moléculas de D-glucosa, conectadas por enlaces glicosídicos α -(1-6), las ramificaciones suelen estar conectadas por enlaces α -(1-3) y contiene azufre de forma heterogénea en las posiciones 2, 4 o 3 de los anillos de la glucosa (figura 4.33). Tres de las cuatro moléculas de SIGN-R1 de la unidad asimétrica mostraron densidad electrónica positiva en el sitio de unión a carbohidratos dependiente de calcio. Además se observó la presencia de monosacáridos en otra zona del receptor, no dependientes de calcio y que en este trabajo lo designaremos como sitio de unión secundario, no canónico. En la figura 4.33 se esquematiza la estructura de un fragmento de sulfodextrano indicando en recuadros las fracciones observadas en los complejos.

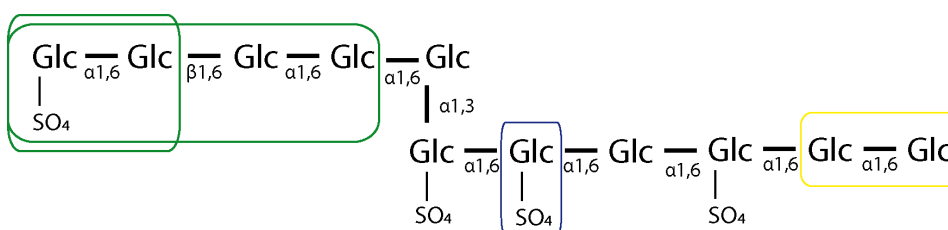


Figura 4.33. Representación esquemática del sulfodextrano. Se indica en cuadros de colores las porciones del polímero observadas en las estructuras obtenidas. En el sitio primario dependiente de calcio, se observaron los fragmentos, disacárido sin azufre en amarillo, tetrasacárido y disacárido sulfatado en verde. En el sitio secundario no canónico se localizaron monosacáridos sulfatados.

El sitio primario de unión a carbohidratos se sitúa sobre el átomo de calcio y se extiende sobre las hebras β 4, β 3 y hacia la hélice α 3. Comenzaremos definiendo el sitio de unión para el disacárido D-glc- α -(1-6)-D-glc de sulfodextrano presente en la cadena B, el cual no posee sustituciones de sulfatos. En este complejo, el anillo primario Glc1 está estabilizado a través de una red de puentes de hidrógeno con el calcio y los aminoácidos que coordinan este calcio, a través de el grupo hidroxilo del carbono 4 (OH-4) del anillo de la glucosa, interactuando directamente con el átomo de calcio (2.5 Å), con el Glu255 (2.7 Å) y con la Asn257 (3.2 Å). El grupo hidroxilo OH-3 está estabilizado por Asn257

(3.4 Å) y por la Asn258 (3.2 Å). La Asn258 será uno de los residuos claves en la especificidad y versatilidad del sitio primario de unión a carbohidratos en SIGN-R1, debido a que, en las estructuras que presentan sus homólogos este residuo se encuentra comprometido con los calcio auxiliares, lo que no ocurre en SIGN-R1 como ya se describió anteriormente. El segundo anillo de glucosa Glc2 está estabilizado por la Ser275 (2.6 Å) y Glu266 (2.4 Å) que estabilizan al grupo hidroxilo OH-2 de la glucosa 2 y el Glu266 (3.1 Å y 3.4 Å) interactúa con el grupo hidroxilo del carbono 3 (figura 4.34).

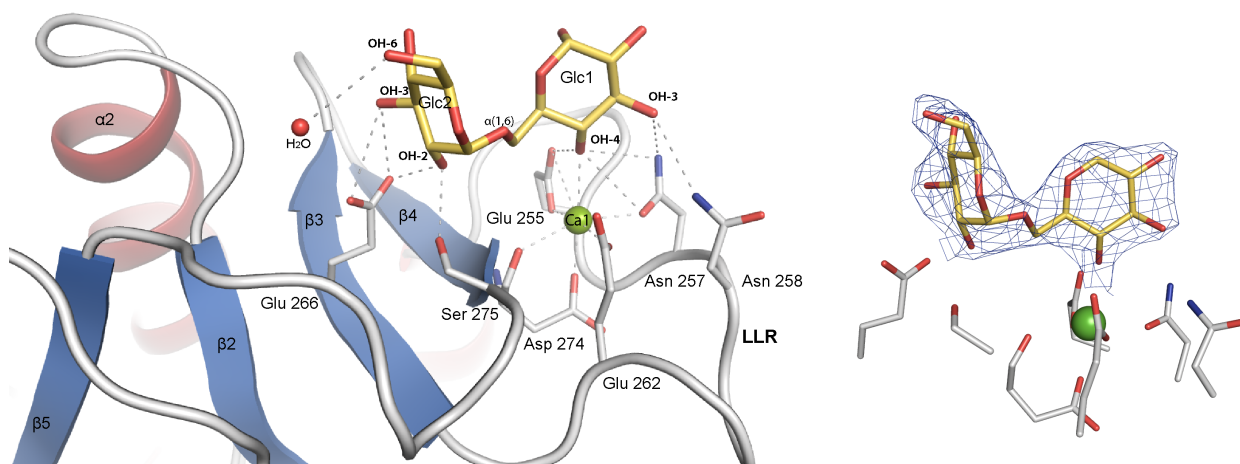


Figura 4.34. Complejo del disacárido de sulfodextrano. El disacárido D-Glc- α -(1-6)-D-Glc se representa en color amarillo, se muestra la numeración de los carbonos de los anillos que interactúan con la proteína. A la derecha se representa la superposición del mapa de densidad electrónica 2Fo-Fc con un contorno de 1.5 σ (azul).

Para la cadena C y D encontramos un disacárido y un tetrasacárido del sulfodextrano respectivamente, D-Glc- α -(1-6)-D-Glc- α -(1-6)-D-Glc- α -(1-6)-D-Glc-SO₄, en la cual se observa una sustitución por sulfato en la posición 4 del anillo de la glucosa en posición 1. Para ambos complejos se observó la misma orientación así como también las interacciones de estabilización, por lo que sólo describiremos el tetrasacarido.

En este complejo, el anillo de glucosa en la posición 1 presenta una orientación distinta en comparación con la Glc1 del disacárido, descrito en el complejo anterior. En el tetrasacárido el oxígeno O1 de la Glc1, está interactuando con el calcio (3.2 Å) y el carbono 4 de este anillo adopta la posición del carbono 1 del disacárido descrito anteriormente (figuras 4.34 y 4.35). En el tetrasacárido, los grupos hidroxilo de los carbonos 1 y 2 son estabilizados por la Asn257 y el Glu255 respectivamente (3.5 Å y 3.2 Å) y el grupo hidroxilo del carbono 3 está formando una nueva interacción con la Asn273 (3.4 Å). El anillo Glc-2 está estabilizado de la misma forma que la Glc-2 para el complejo con el disacárido (figura 4.34), añadiendo una nueva interacción con la Lys219 (figura 4.35). Los anillos Glc-3 y Glc-4 están estabilizados por interacciones hidrofóbicas con la Leu279, Pro221 y Phe282. Además el anillo Glc-4 está estabilizado por el grupo hidroxilo de la cadena principal de la Lys281 (2.4 Å) y por un puente de hidrógeno entre el OH-6 y la Gln183 (3.9 Å) (figura 4.35).

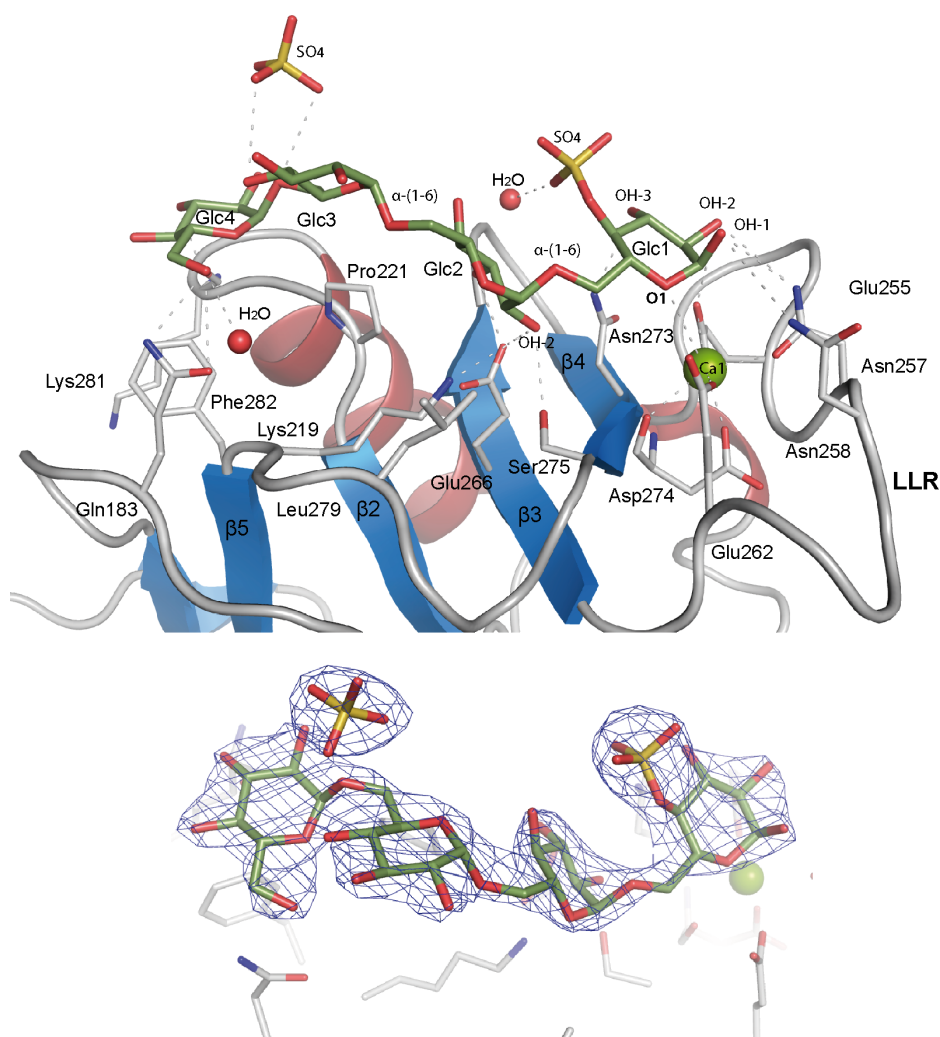


Figura 4.35. Complejo del tetrasacárido de sulfodextrano. El tetrasacárido 2(D-glc- α -(1-6)-D-glc) se representa en color verde, se muestra la numeración de los carbonos de los anillos que interactúan con la proteína. Debajo se representa la superposición del mapa de densidad electrónica 2Fo-Fc con un contorno de 1.5 σ (azul).

El análisis de los mapas de densidad electrónica alrededor de la cadena C en complejo SIGN-R1:Sulfodextrano reveló la presencia de tres moléculas de glucosa-4-sulfato. Estos tres monosacáridos están situados en la cara opuesta a la zona de unión primaria a carbohidratos, ocupando las posiciones de los sulfatos S2-S4 presentes en la estructura nativa de SIGN-R1 (figuras 4.32 y 4.36). En este complejo tanto los anillos de glucosa como los sulfatos están fuertemente estabilizados (figura 4.36). La glucosa 1 (Glc1) sustituye al sulfato S4 y está estabilizado por un puente de hidrógeno (2.60 Å) entre la Lys249 y el sulfato de esta glucosa y otro puente de hidrógeno (3.11 Å) entre el grupo hidroxilo del carbono 2 de la glucosa y el oxígeno carbonilo de la Arg226. Para la glucosa que observamos en la posición central (Glc2) y que sustituye al S3, son la Asn204 y la Arg246 que se encuentran formando puentes de hidrógeno (3.00 Å) con los grupos hidroxilos de los carbonos 2 y 1 de la Glc2 respectivamente, y el sulfato de esta glucosa está estabilizado por los grupos aminos la Asp206 y Glu207 de la cadena principal (3.10 Å y 3.05 Å). La Glc3, está estabilizada por puentes de hidrógeno entre Asn175 y el grupo hidroxilo del carbono 3 (3.2 Å) y Lys286 con el grupo hidroxilo del carbono 2 (2.6 Å), y el sulfato de esta glucosa está estabilizado por los residuos Asn175, Ser205 y Glu208 (3.2 Å, (3.3 Å y 2.6 Å) (figura 4.36).

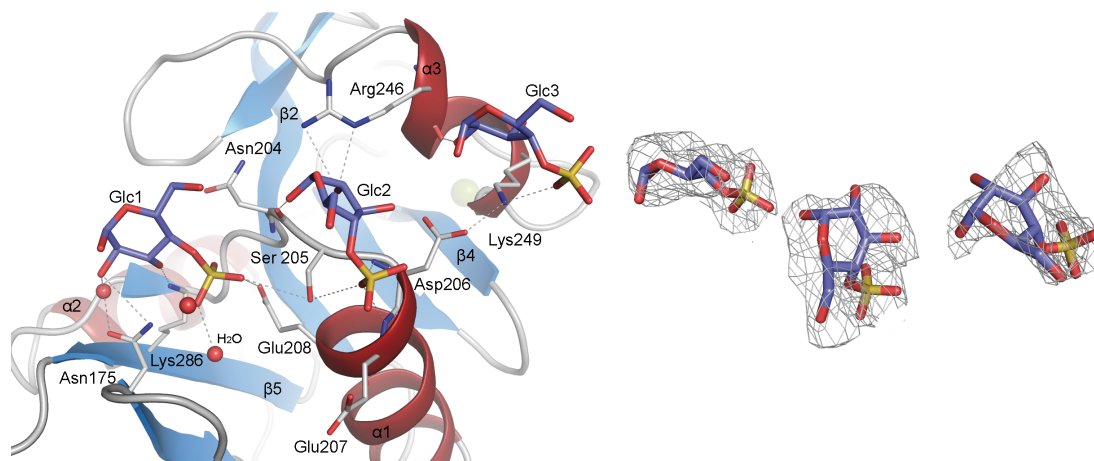


Figura 4.36. Complejo de monosacáridos de sulfodextrano. Se muestra las glucosas-4-sulfato y los aminoácidos e interacciones que estabilizan el complejo. A la derecha se representa la superposición del mapa de densidad electrónica 2Fo-Fc con un contorno de 1.2 σ (gris).

4.2.6.2 Complejo CRD_SIGN-R1 : Ácido siálico

Para las cuatro cadenas de la unidad asimétrica, el ácido siálico se observa en la misma posición y similar orientación, por eso analizaremos en detalle una de ellas. La molécula de ácido siálico está estabilizada a través de una red de puentes de hidrógeno entre, el átomo de calcio Ca1 y el ácido carboxílico del ácido siálico, este grupo sustituye dos de los sitios de la coordinación octaédrica del calcio a 2.3 Å y 3 Å (figura 4.37). A su vez estos dos grupos carboxílicos del ácido siálico están siendo estabilizados por la Asn257 (2.5 Å) y Asn258 (2.8 Å) y además por Glu255 (3.1 Å). Luego tenemos una molécula de agua en medio del calcio y una molécula de sulfato que a su vez está estabilizada por la Ser275.

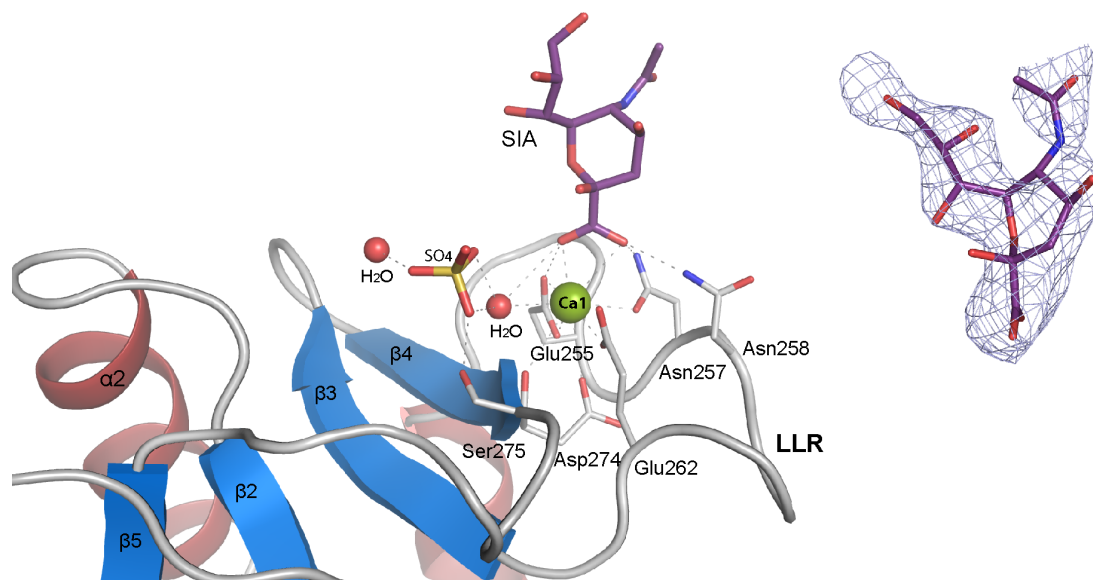


Figura 4.37. Complejo con ácido siálico. El ácido siálico se representa en color morado, el átomo de calcio se representa como una esfera de color verde, se muestra también. A la derecha se representa la superposición del mapa de densidad electrónica 2Fo-Fc con un contorno de 1 σ (azul).

4.2.7 Cálculos de “docking” con ribitol fosfato y cápsula polisacárida

La disposición repetitiva de los moléculas de sulfato reconocidas por SIGN-R1 a través del sitio de unión secundario nos indujo a explorar el posible reconocimiento de este receptor por otros componentes de superficie bacteriana con estructuras repetitivas. El ácido teicoico es el componente más abundante de la pared celular de bacterias Gram positivas y además está formado por cadenas donde la unidad repetitiva es azúcar-fosfato. Estos glicopolímeros frecuentemente están formados por grupos de glicerol o ribitol unidos entre sí a través de enlaces fosfodiéster, como ocurre por ejemplo en *Staphylococcus aureus*, *Bacillus subtilis* o *Listeria monocytogenes* (Weidenmaier & Peschel, 2008; Brown *et al.*, 2010). Por esta razón decidimos realizar un estudio de “docking” con un fragmento de poliribitol-fosfato (pRF) que contenía cuatro unidades. Los cálculos resultantes arrojaron 10 soluciones muy cercanas entre sí, donde los átomos de fosfato tomaron los lugares de los sulfatos S2-S4 reconocidos por SIGN-R1 en el sitio de unión 2 (figura 4.38). Este sitio presenta un gran surco de 30 Å de longitud aproximadamente en donde predominan residuos básicos, formado en casi su totalidad por el bucle que conecta las hélices alfa $\alpha 2$ y $\alpha 3$, por la hélice $\alpha 3$ y la lamina $\beta 2'$.

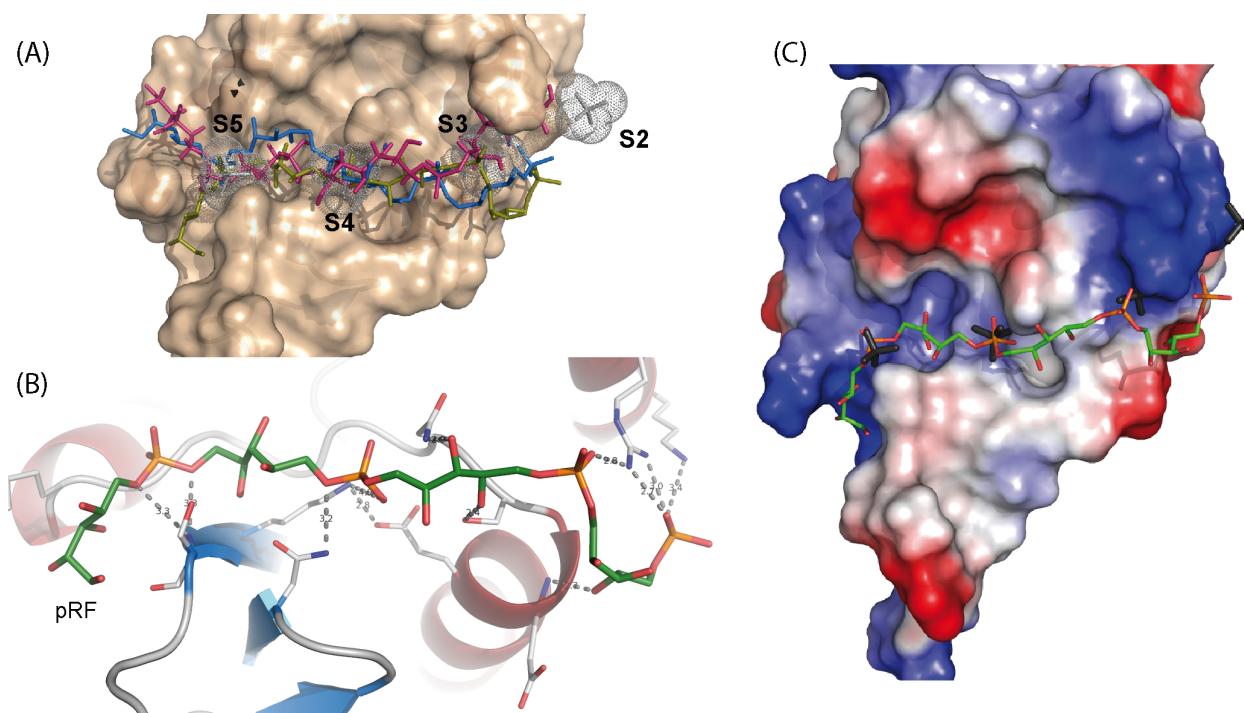


Figura 4.38. Resultados de los cálculos del “docking” con poliribitol fosfato en el sitio de unión 2. (A) Se muestran tres soluciones calculadas y superposición con los sulfatos de la estructura nativa. (B) Detalle de las interacciones. (C) Superficie de potencial electrostático. pRF, poliribitol fosfato (verde), S2-S5, sulfatos (negro).

De igual forma se realizó un ensayo de “docking” con un fragmento de la CPS14, ligando natural de SIGN-R1, para verificar si este polisacárido, podría ser acomodado en el sitio secundario de unión. Los resultados obtenidos indican que la CPS14 podría ser reconocida por el sitio secundario (figura 4.39). Además los anillos de los azúcares de la CPS14 reemplazan a las glucosas obtenidas en el complejo de SIGN-R1:Sulfodextrano.

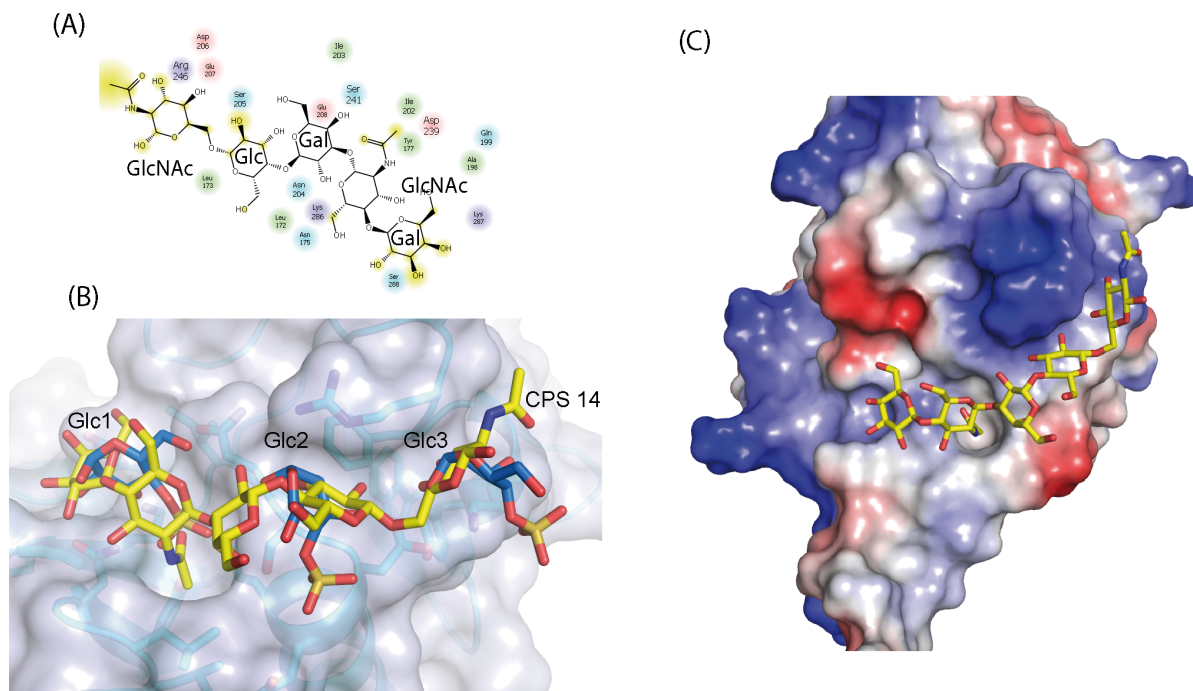


Figura 4.39. Resultados del cálculo de “docking” para el el pentasacárido de la CPS14. (A) Diagrama de los aminoácidos y sus interacciones implicados en el reconocimiento de la CPS14 y formación de la cavidad. (B) Superposición con el complejo cristalográfico SIGN-R1:Sulfodextrano. (C) Superficie de potencial electrostático del complejo por “docking” de CPS 14. CPS14, cápsula polisacárida de neumococo del serotipo 14 (amarillo). Glc1-Glc3, glucosas-4-sulfato, azul.

4.3 DETERMINACIÓN ESTRUCTURAL DE LA PORCIÓN Fc

4.3.1 Cristalización y optimización

Con el objetivo de determinar la estructura tridimensional de la porción Fc con cadenas glicánicas sialiladas, se procedió a cristalizar muestra de este anticuerpo producidos en el laboratorio del Dr. Robert M. Anthony del laboratorio de genética molecular e inmunología de la Universidad Rockefeller (New York).

Las condiciones iniciales de cristalaron se llevaron a cabo bajo técnicas de alto rendimiento. Al cabo de 10 días se comenzó a observar la formación de material cristalino en forma de placas de fino espesor, agujas y pseudo-cristales con alta nucleación, como puede observarse algunos ejemplos en la figura 4.40-A (fotografías 1, 2, 3 y 4). Estas condiciones tenían en común la naturaleza del agente precipitante, polietilenglicol (PEG), en muchos de los casos cloruro o formiato sódico como aditivo y diferían en el pH y tipo de tampón. La estrategia a seguir para el escalado fue hacer un barrido de las concentraciones de precipitante a diferentes pH y en ausencia/presencia de aditivo. El resultado de este escalado fue la obtención de cristales con alto grado de macla (figura 4.40-B fotografías 5 y 6). Para resolver este problema se ensayaron variaciones de gota. Sólo se obtuvieron monocristales en las proporciones 1:3 (1 μ l de proteína : 3 μ l de reservorio), los cuales alcanzaron un tamaño de 0.7 x 0.3 x 0.4 mm en tres semanas (figura 4.40-C, fotografía 7).

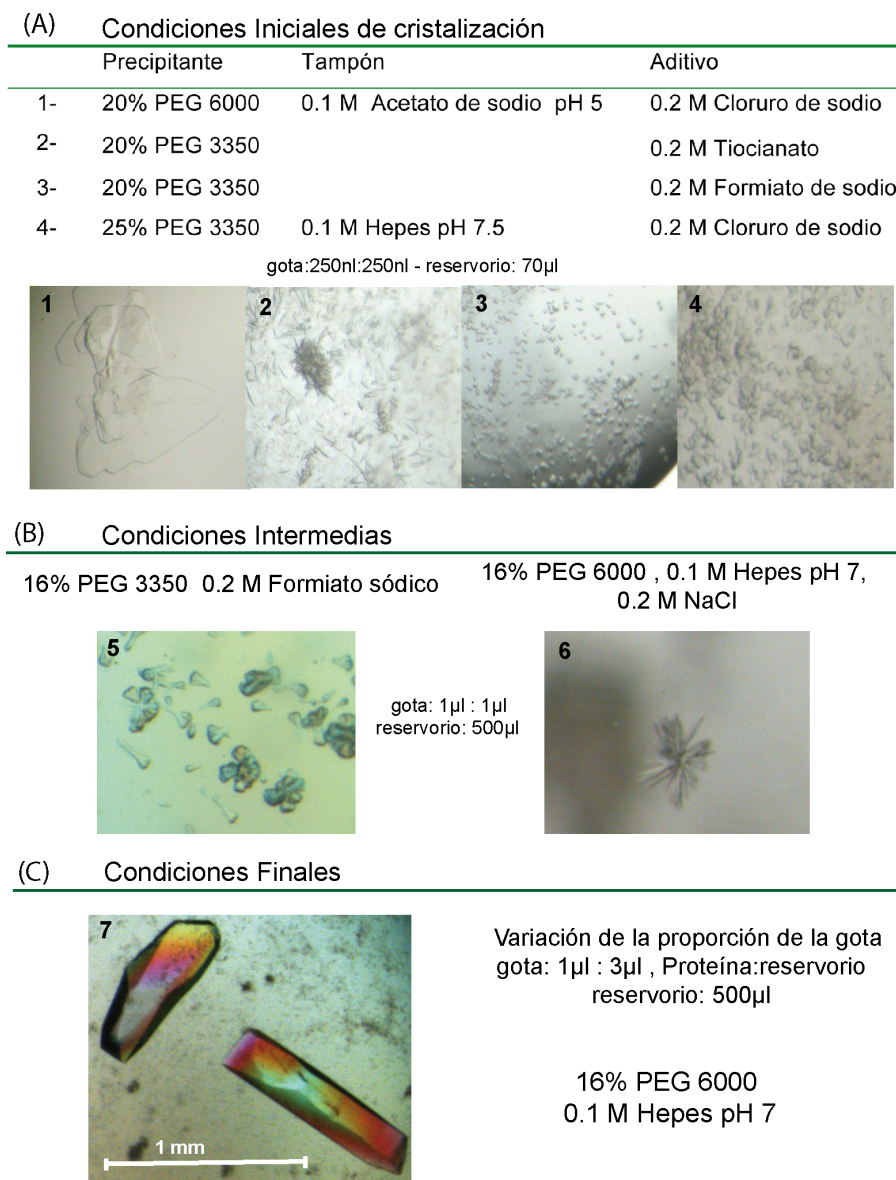


Figura 4.40. Cristalización del Fc_Sial. (A) Ensayos iniciales de cristalización con técnicas de alto rendimiento, cuyas condiciones derivan de los siguientes screening 1- PACT-A7, 2- PACT-E4, 3- PACT-E6 y 4- Index-F12. Las 4 primeras fotos corresponden a dichas soluciones comerciales de cristalización. (B) Escalado de las condiciones iniciales. (C) Condición final de cristalización.

4.3.2 Medida de los datos de difracción

Los cristales de mayor tamaño fueron crio-protegidos en una solución que contenía 25% de glicerol en solución madre. Se midió un conjunto de datos a 2.01 Å de resolución, obteniendo patrones de difracción de excelente calidad (figura 4.41) recogidos en la línea ID14-4 del ESRF. El resultado del indexado indicó que los cristales pertenecían al sistema cristalino ortorrómbico $P2_12_12_1$, presentando unas dimensiones de celdilla de $a = 49.96$ Å, $b = 80.75$ Å y $c = 138.40$ Å. Las estadísticas del procesado y escalado de los datos de difracción se listan en la tabla de la figura 4.41.

Teniendo en cuenta la masa molecular del Fc_Sial y el volumen de la celdilla unidad, obtuvimos un coeficiente de Matthews de 2.74 Å³·Da⁻¹ (Matthews, 1968) para el dímero biológico en la unidad asimétrica con un contenido de solvente del 55.06%.

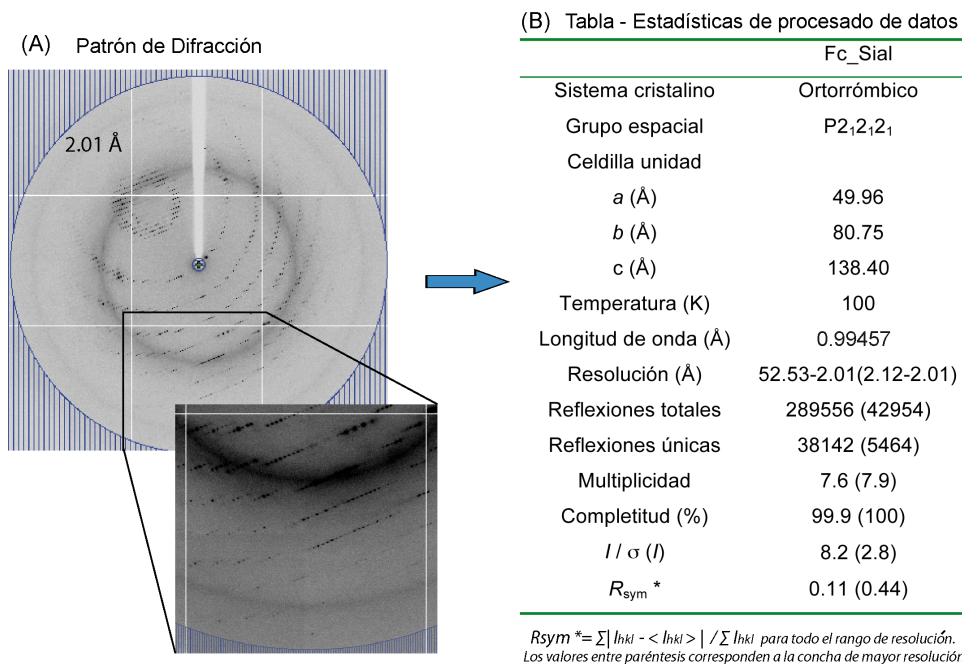


Figura 4.41. Datos de difracción de Fc_Sial. (A) Patrón de difracción a 1° grado de oscilación. En el recuadro aumentado se muestra reflexiones a 2.01 Å de resolución. (B) Estadística del procesamiento.

4.3.3 Resolución estructural y refinamiento

La resolución estructural se llevó a cabo mediante el Método de Reemplazo Molecular, utilizando como modelo estructural el monómero de la fracción Fc del IgG1 humano (código PDB: 1HZH) que presenta un 99% de identidad de secuencia para esta fracción del anticuerpo. En el Reemplazo Molecular se obtuvieron dos soluciones contrastadas del resto. La correlación fue de 25,2% y el factor R de 54.8%. Esta baja correlación puede deberse al hecho de que las cadenas polipeptídicas del fragmento Fc, son extremadamente flexibles a nivel “intra” e “inter” dominio a pesar de tener un alto grado de homología entre los diferentes tipos de IgG e inclusive entre diferentes especies (Saphire *et al.*, 2002).

El mapa de densidad electrónica calculado nos permitió modelar y ajustar perfectamente cada cadena polipeptídica de 205 aminoácidos para la cadena A y 207 para la cadena B (residuos 236 - 244), una cadena glicánica con ramificación doble de 10 unidades de azúcar para la cadena polipeptídica A y 8 unidades de azúcar para la cadena B. Una vez refinada la estructura, se comprobó la estereoquímica a través del Diagrama de Ramachandran. El 92.58% de los residuos se encuentran dentro de las regiones favorables, el 7.42% de los residuos se observaron en regiones permitidas y ninguno en las zonas prohibidas. Las estadísticas finales del refinamiento y el diagrama de Ramachandran se muestran en la figura 4.42.

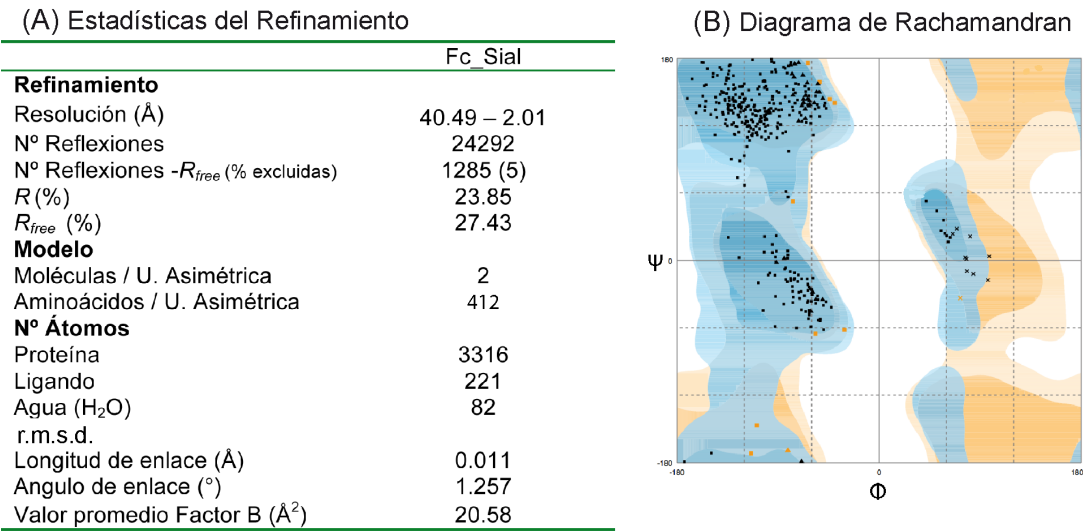


Figura 4.42. Refinamiento del modelo de la estructura de Fc_Sial. (A) Estadísticas del refinamiento de Fc_Sial. (B) Diagrama de Ramachandran, los cuadros muestra los residuos y la diferencia de color negros y amarillos muestran si son favorables y permitidos respectivamente.

4.3.4 Análisis de la estructura tridimensional

4.3.4.1 Descripción general

La estructura tridimensional de los IgGs, ha sido ampliamente estudiada, están constituidos por tres dominios independientes (cadenas pesadas y ligeras) conectados por regiones flexibles. Dos de estos dominios están estructuralmente englobados en el llamado fragmento Fab, el cual tiene la especificidad para la unión de los antígenos. El tercer dominio estructural está asignado al fragmento Fc, el cual tiene la propiedad de ser reconocido por receptores, activando o desactivando diversos mecanismos (Huber *et al.*, 1976) (figura 4.43-A).

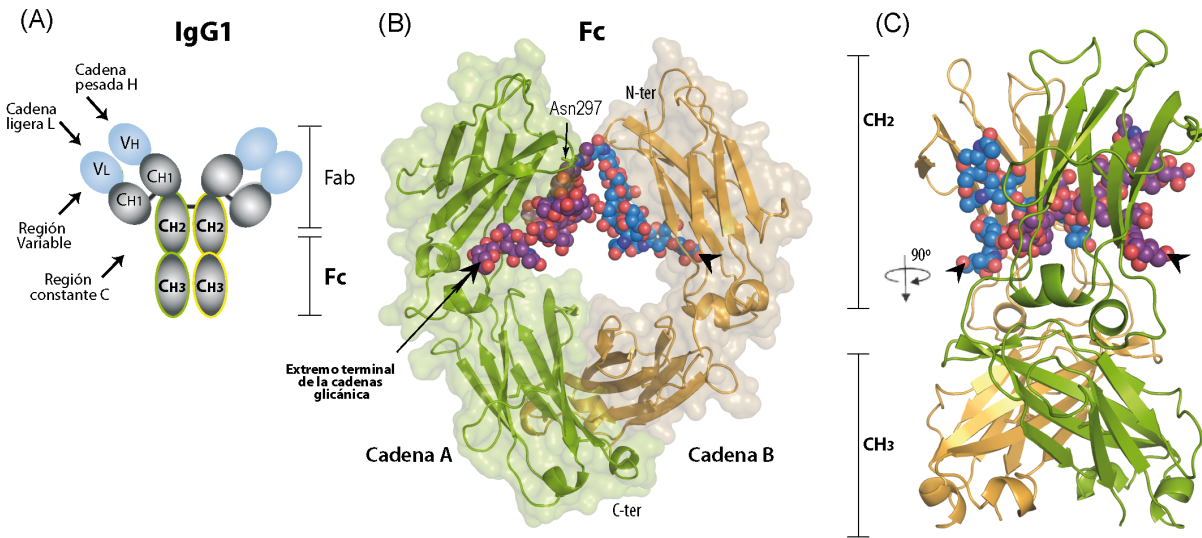


Figura 4.43. Estructura tridimensional del fragmento Fc_Sial. (A) Esquema general de la organización estructural del IgG. El IgG se compone del fragmento Fab y del fragmento Fc. El fragmento Fab está constituido por la región variable (VH y VL) y la región constante 1 (CH1). El fragmento Fc está formado por la región constante 2 y 3 (CH2 y CH3). (B) Estructura cristalográfica del Fc_Sial, homodímero formado por las cadenas A (verde) y B (amarillo), divididas a su vez por los dominios CH2 y CH3. Dentro de la estructura en forma de herradura se muestran las glicosidaciones, en morado para la cadena A y en azul para la cadena B. Se señala la Asn297 de la cadena A de donde comienza la glicosidación, (C) Rotación de 90° grados para la estructura general, donde se observan las cadenas glicánicas en el interior, dejando expuesto sólo los azúcares terminales señalados con una punta de flecha.

La estructura tridimensional del Fc tiene forma de herradura y consta de dos cadenas polipeptídicas A y B (homodímero), formado por los dominios de la cadena pesada 2 y 3 (CH2 y CH3) de la inmunoglobulina, (figura 4.43-B). La determinación estructural del fragmento Fc presentó una glicosidación en ambas cadenas, ubicada en el extremo N-terminal del dominio CH2 en la Asn297, esta glicosidación comienza en la zona interna o central de la herradura y se extiende hacia las zonas externas de la herradura por ambos lados, en donde sólo los azúcares terminales quedan expuestos, figura 4.43-B y 4.43-C.

4.3.4.2 Análisis estructural de las cadenas glicánicas

Cada una de las cadenas del dímero tiene una N-glicosilación en la Asn297. Cada oligosacárido observado está constituido por una cadena ramificada, de la cual se han podido determinar perfectamente 10 unidades de azúcar para la cadena A, una fucosa (Fuc; posición 2), cuatro N-acetilglucosaminas (GlcNAc; posición 1,3,6 y 9), tres manosas (Man; posición 4,5 y 8) y las dos galactosas terminales (Gal; posición 7 y 10), mientras que para la cadena B se han observado 8 unidades de azúcar, cuatro GlcNAc (posición 1,3,6 y 9), tres manosas (posición 4,5 y 8) y una galactosa terminal (posición 7). En la figura 4.44-A se muestra esquematizada la estructura de la cadena glicánica, así como la estructura tridimensional determinada en la cadena A (figura 4.44-B).

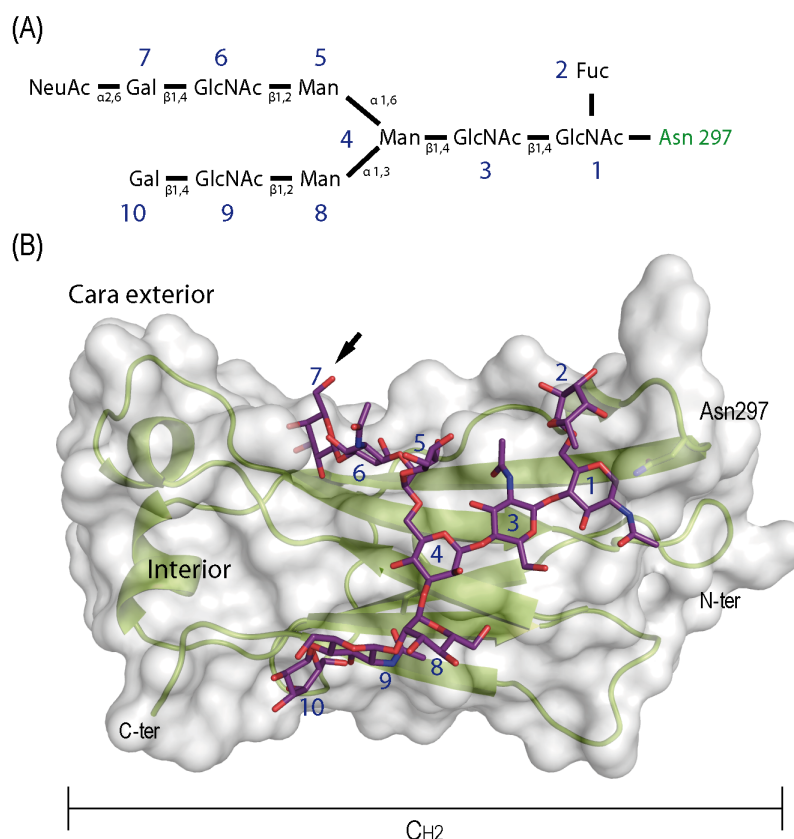


Figura 4.44. Estructura tridimensional de la cadena glicánica observada (A) Representación esquemática de la cadena glicánica ramificada, se muestra la numeración para las posiciones de cada unidad de azúcar. (B) Estructura tridimensional de la glicosidación para la cadena A, se muestra la N-glicosilación en la Asp297 y GlcNAc 1. Con una flecha se indica el grupo hidroxilo del carbono 6 de la galactosa terminal.

La cadena glicánica está fuertemente estabilizada por una red de puentes de hidrógeno e interacciones hidrofóbicas a lo largo de toda su estructura, y termina en este caso en la galactosa que

hemos determinado como posición 7 y 10, la cual deja expuesto un grupo hidroxilo del carbono 6 de este azúcar. De esta forma a pesar de que en esta estructura no se observó el ácido siálico, la galactosa terminal tiene una conformación favorable y disponible para el enlace glicosídico α -2,6 entre la galactosa y el ácido siálico (figura 4.44).

Para confirmar la orientación del ácido siálico terminal, con un enlace glicosídico α -2,6, se realizaron ensayos de dinámica molecular para el residuo de ácido siálico (NeuAc) al Gal terminal número 7 y 10. Se observó en primer lugar que el ácido siálico puede tener una gran movilidad y segundo que dejaría expuesto al grupo carboxílico (figura 4.45-A). Cabe la pena destacar la orientación estructural del ácido siálico en posición α -2,3 en una estructura homóloga ya publicada, la cual es de la única que se dispone información estructural de la cadena glicánica hasta el ácido siálico (Código PDB: 1MCO). En esta estructura, el ácido siálico ya no tiene la misma movilidad debido a que el enlace glicosídico α -2,3 es más corto de lo que puede llegar a ser un enlace glicosídico α -2,6 y el grupo carboxílico del NeuAc está enterrado en la estructura polipeptídica (figura 4.45-B).

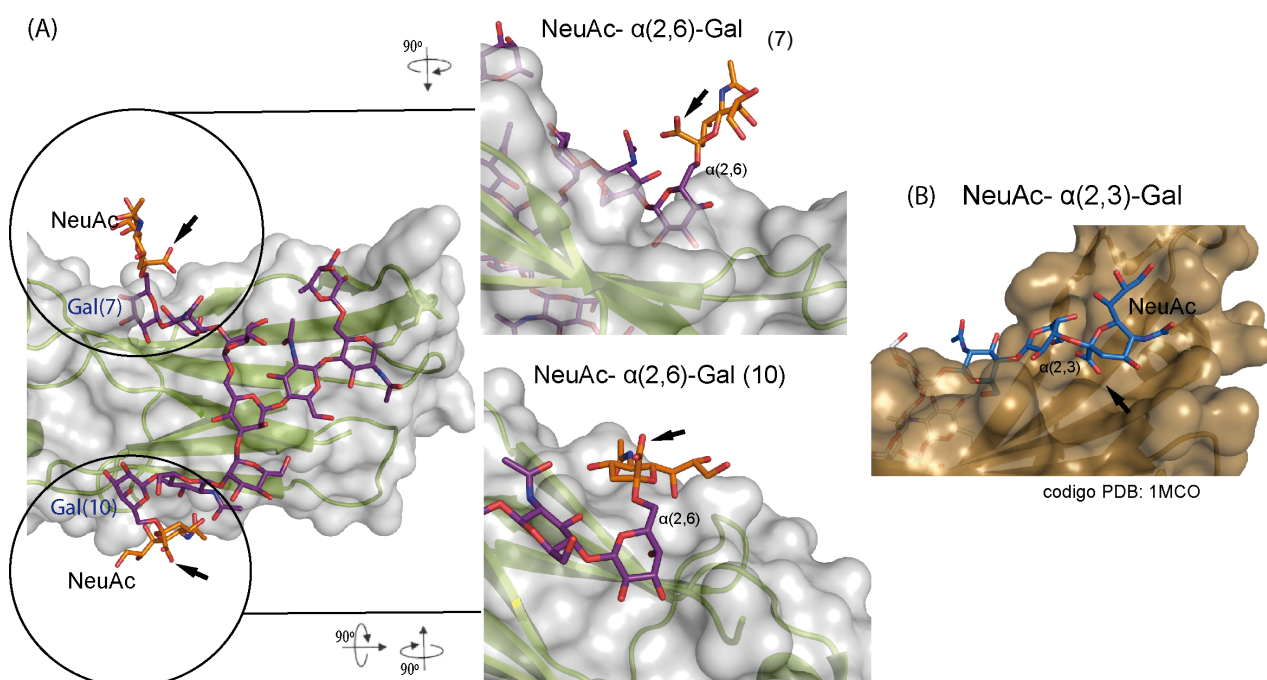


Figura 3.45. Ensayos de dinámica molecular para el ácido siálico terminal unido a la galactosa 7 y 10. (A) Resultados de dinámica molecular para el enlace glicosídico α -2,6 entre la Gal (7) arriba y la Gal (10) abajo. (B) Estructura tridimensional del Fc sialilado en α -2,3 (código PDB:1MCO). Los grupos carboxílicos de los ácidos siálicos en todos los casos se señalan con una flecha.

DISCUSIÓN

5 DISCUSIÓN

5.1 ESTRUCTURA DE LA ENDOLISINA Cpl-7

Todas las enzimas líticas de neumococo caracterizadas hasta el momento, ya sean codificadas por la bacteria o por los bacteriófagos que la infectan, presentan una organización modular que podría resultar de una adaptación evolutiva dirigida al reconocimiento de motivos moleculares de la pared bacteriana. La adquisición de un módulo de anclaje al sustrato en estas enzimas hidrolíticas se traduciría en un aumento de la eficacia catalítica (Mitchell *et al.*, 1997; Tuomanen, 1999; Berry & Paton, 2000).

Aunque la mayoría de las enzimas que hidrolizan la pared de neumococo presentan un módulo de unión a la pared especializado en reconocer las moléculas de colina que define a la familia de las CBPs, la endolisina Cpl-7 es la excepción. Se ha demostrado la efectividad de las endolisinas Pal y Cpl-1, como agentes antimicrobianos en el tratamiento de infecciones tanto locales como generalizadas, y la capacidad de las mismas para actuar sinérgicamente entre sí o en combinación con antibióticos clásicos como la penicilina y la gentamicina (Jado *et al.*, 2003; Djurkovic *et al.*, 2005; McCullers *et al.*, 2007; Rodríguez-Cerrato *et al.*, 2007a). Sin embargo, la desventaja que presentan tanto Pal como Cpl-1 es que son enzimas que pertenecen a la familia de las CBPs, por lo que su actividad como enzibióticos está limitada a microorganismos que contengan colina en su envuelta celular.

Este hecho abre nuevas posibilidades en el uso potencial de Cpl-7 como enzibiótico ideal y universal, ya que el módulo de unión a la pared se ha identificado en proteínas implicadas en la degradación del peptidoglicano en otras bacterias patógenas, como es el caso de la endolisina del profago λ Sa2 de *S. agalactiae*. Esta endolisina es capaz de lisar varias especies (Donovan & Foster-Frey, 2008). Por lo tanto Cpl-7 podría ser activa frente a diversos patógenos del género *Streptococcus*.

5.1.1 Producción de Cpl-7 como proteína de fusión con la lectina LSL y cristalización

El primer objetivo de este trabajo fue la puesta a punto de un método que permitiera purificar a homogeneidad y en grandes cantidades la proteína Cpl-7, mediante un procedimiento relativamente sencillo y en pocos pasos que incluye dos cromatografías, de afinidad y exclusión molecular (paso intermedio, corte con TEV). Se ha conseguido purificar a homogeneidad y con un buen rendimiento la lisozima Cpl-7. Con esta muestra purificada y homogénea de Cpl-7 se consiguieron cristales de excelente calidad. La muestra proteica sufrió degradación en los ensayos de cristalización a pesar de que las muestras estaban suplementadas con inhibidores de proteasas y azida sódica para evitar contaminaciones. La degradación dio lugar a la cristalización del módulo catalítico de Cpl-7. No se pudieron obtener cristales de la proteína completa a pesar de todos los ensayos realizados con: técnicas convencionales, el sembrado de cristales, uso de aditivos y detergentes, adición de posibles sustratos (fragmentos de peptidoglicano purificado de neumococo y fragmentos de ADN de longitudes

variables) y ligeras modificaciones en el proceso de purificación como el cambio de sales y solución tampón.

Por estas razones emprendimos la obtención del módulo de unión a la pared, utilizando una estrategia de producción similar a la seguida para la obtención de Cpl-7 completa. Este método nos permitió purificar a homogeneidad en altas concentración dicho módulo. En los cristales nos enfrentamos nuevamente con la proteólisis de las repeticiones a través de su conector, ya que el contenido de los péptidos en los cristales arrojaron masas moleculares acordes con una y dos repeticiones (figura 4.14). Se descarta la posibilidad de corte con la proteasa TEV ya que se analizó la secuencia de Cpl-7 y no posee secuencia de corte frente a esta proteasa altamente específica.

5.1.2 El módulo catalítico de Cpl-7 sigue el plegamiento de las GH-25

El módulo catalítico de Cpl-7 pertenece a la familia GH-25 de las glicosilhidrolasas, y presenta una estructura de tipo barril α/β irregular, $(\alpha/\beta)_5 \beta_3$. El sitio catalítico está situado en la cara del barril en la que se encuentra el extremo C-terminal de las hebras β (figura 4.7-A y 4.7-B). Hacia el centro de la cavidad catalítica se observa la inusual presencia de una hebra β antiparalela (β_8) que determina la posición de uno de los residuos implicados en la actividad. Estas características estructurales aparecen en todas las lisozimas pertenecientes a esta familia, tales como: en bacterias, lisozima de *Bacillus anthracis* (código PDB: 2WAG), cellosyl, una autolisina de *Streptomyces coelicolor* (código PDB: 1JFX) y LytC, una autolisina de *S. pneumoniae* (código PDB: 2WW5), en bacteriófagos, Cpl-1 (código PDB: 1HO9) y plyB (código PDB: 2NW0) y en hongos, la lisozima de *Aspergillus fumigatus* (código PDB: 28XR).

La cavidad catalítica de Cpl-7 presenta dos pares de residuos ácidos y enfrentados (E94-D92 y D10-D182) (figura 4.7-B), residuos que se encuentran conservados en todos los miembros de la familia. Este par de residuos juegan un papel crucial en la catálisis enzimática, como se deduce de la drástica pérdida de la actividad que supone la mutación a alanina o glutamina en los residuos equivalentes en Cpl-1 (Hermoso *et al.*, 2003). La disposición de este par de residuos definidos en el módulo catalítico y la distancia específica entre ellos (8.43 Å), indicaría que la hidrólisis catalizada por Cpl-7 seguiría un mecanismo con inversión de la configuración del carbono anomérico, al igual que lo hacen las lisozimas de la familia GH-25. Así, en Cpl-7, el residuo Glu94 actuaría protonando por la cara β al oxígeno glicosídico del enlace entre los azúcares -1 y +1, mientras que el residuo Asp10 activaría una molécula de agua que se interpone entre él y la cadena glicánica para que efectúe el ataque nucleofílico, invirtiendo la posición de C1 del azúcar a la forma α , como se describió para la lisozima del fago T4 (figura 1.14-B).

5.1.3 Reconocimiento del sustrato por el módulo catalítico de Cpl-7

La cavidad catalítica de Cpl-7 es profunda y presenta además un surco lateral. La interacción de los residuos que forman este surco con el sustrato podría determinar la orientación de la cadena glicánica y del enlace a hidrolizar dentro del centro catalítico (figura 4.7-C y 4.8-B). Esta forma de

túnel se ha asociado a enzimas hidrolíticas que efectúan un ataque múltiple en diversos puntos de la pared sin liberarse del sustrato (Henrissat & Davies, 1997). La similitud observada en la superposición de las estructuras cristalografías de los módulos catalíticos de Cpl-7 y Cpl-1 (r.m.s.d = 0.36 Å, figura 4.8-A) indica, que en Cpl-7 la interacción con el sustrato será equivalente a lo reportado para Cpl-1. Así en Cpl-7 la cavidad catalítica está formada por las cadenas laterales de residuos aromáticos conservados y expuesto a la superficie (Y59, F61, Y125, Y127, P129, W149 y Y153 figura 4.9). Estos residuos podrían participar en la unión mediante apilamiento de los anillos aromáticos con la cara polar de los azúcares de la cadena glicánica como se ha descrito en Cpl-1 y en otras glicosilhidrolasas y en proteínas que unen carbohidratos (Nagy *et al.*, 1998; Simpson & Barras, 1999; Hermoso *et al.*, 2003; Pérez-Dorado *et al.*, 2007; Pérez-Dorado *et al.*, 2010).

Particularmente interesante es el residuo Tyr127 que parece jugar un papel de tapadera (*gate keeper*) del sitio activo. En efecto, este mismo residuo se ha observado en Cpl-1 en dos conformaciones diferentes: cerrada (en la estructura cristalográfica nativa), bloqueando el acceso al sitio catalítico y abierta, permitiendo el acercamiento del sustrato a los residuos catalíticos y estabilizando el anillo de MurNAc en posición +1 (Pérez-Dorado *et al.*, 2007). En el caso de Cpl-7, la Tyr127 se encuentra en una conformación cerrada, es por tanto de esperar que juegue un papel similar en su conformación abierta en la estabilización del sustrato previo a la hidrólisis.

En Cpl-1 se ha identificado una red iónica próxima al par donador de protones formada por los residuos (K34, E37, R63, D95 y D96). Esta red iónica participaría en la unión del sustrato formando puentes de hidrógeno, como indican los cambios de actividad asociados a la sustitución de E37 por otros aminoácidos polares o alanina (Sanz *et al.*, 1992b), además estos aminoácidos están implicados directamente en la estabilización de la porción peptídica del MurNAc en posición -1 (Pérez-Dorado *et al.*, 2007). En el caso particular de Cpl-7 observamos dos mutaciones con respecto a Cpl-1 (R63W y D96H), (figura 4.9-B), que podrían afectar el reconocimiento al sustrato, proporcionando una mayor estabilidad para los anillos de azúcares por parte del triptófano y una mayor estabilidad para la porción peptídica del PG que estén cargados negativamente, como el extremo C-terminal de algunos muropéptidos, o interaccionando con otra cadena glicánica adyacente. Otra sustitución importante dentro de la cavidad catalítica de Cpl-7 es la Ala12 (serina en Cpl-1), la cual sería la responsable de la estabilización del anillo de azúcar en la posición -1, esto podría reducir el número de puentes de hidrógeno con el sustrato y provocar cambios en la actividad de Cpl-7.

En cuanto al resto de mutaciones no conservativas presentes en el módulo catalítico de Cpl-7 con respecto de Cpl-1 (V68E; A69E, M26A, M84T y H131L), se sitúan, mayoritariamente, fuera de la cavidad catalítica. Estas mutaciones no suponen cambios estructurales y por tanto no afectarían, *a priori*, a la interacción con el sustrato. Vale la pena destacar la presencia del canal hidrofóbico en Cpl-7, responsable de la estabilización de los módulos catalítico y de unión a la pared en Cpl-1. De todas las estructuras estudiadas de los miembros de la familia GH-25, este canal hidrofóbico sólo se había observado hasta el momento en Cpl-1. Esta observación indicaría la posibilidad de una disposición modular en Cpl-7 similar a la disposición modular en Cpl-1.

5.1.4 Módulo de unión a la pared de Cpl-7

En Cpl-7 el módulo de unión a la pared ocupa la posición C-terminal, y está formado por tres repeticiones idénticas, excepto la última repetición que es más corta ya que carece del conector entre las repeticiones (figura 4.10). Los cristales obtenidos nos permitieron determinar la estructura de una repetición del CWBM de Cpl-7 a 1.67 Å de resolución. La estructura estudiada consta de tres hélices- α , dos de ellas se posicionan de forma paralela y la tercera, más larga, se dispone perpendicular a las dos primeras, también podemos observar la región que conectaría una repetición con la siguiente (figura 4.16 y 4.17). A partir del estudio de la disposición de las repeticiones del CWBM de Cpl-7 en el empaquetamiento cristalino, ha sido posible inferir información sobre la posible disposición *in vivo* de dichas repeticiones en el módulo de anclaje de Cpl-7. El cristal está formado por péptidos constituidos por una y dos repeticiones del CWBM de Cpl-7 tal como se evidenció en el perfil de espectrometría de masas (figura 4.14). Este hecho, unido a la completa identidad entre las repeticiones del CWBM de Cpl-7, provocó la presencia de densidad electrónica perfectamente definida que conecta las repeticiones entre sí. Además se observaron interacciones entre repeticiones que podrían configurar la disposición global del módulo de anclaje de Cpl-7 (figura 4.20).

Recientemente (Bustamante *et al.*, 2010) se ha propuesto un modelo global del módulo de unión de Cpl-7 basado en la estructura del regulador genético RecX (Ragone *et al.*, 2008) que presenta tres dominios homólogos. La superposición del dominio de una repetición del CWBM de Cpl-7 determinado por cristalografía, con la estructura de RecX, indica que sólo existe homología estructural con la repetición central de RecX (figura 4.18). Los aminoácidos que componen el centro hidrofóbico de ese dominio también están conservados entre la repetición central de RecX y las repeticiones de Cpl-7. En contraste el conector entre las repeticiones no está conservado: en RecX son de 3 y 2 aminoácidos (entre R1 y R2) mientras que en Cpl-7 esta región tiene siempre 6 residuos. RecX tiene una forma alargada y su primera y tercera repeticiones no siguen el mismo patrón estructural que la repetición central. Las implicaciones de estas diferencias se detallarán posteriormente.

5.1.5 Especificidad de Cpl-7 por su sustrato

A la hora de considerar la especificidad de las distintas mureín-hidrolasas se debe tener en cuenta que el peptidoglicano bacteriano está altamente conservado entre bacterias a nivel de la cadena glicánica (disacáridos de MurNAc y GlcNAc unidos por enlaces glicosídicos β -1,4) y composición peptídica. Modificaciones en los anillos glicánicos, como la O-acetilación en el C6-OH de MurNAc pueden afectar la capacidad de la lisozima para digerir las cadenas del peptidoglicano (Bera *et al.*, 2006) y sólo unas pocas enzimas tienen la capacidad para hidrolizar cadenas glicánicas con MurNAc O-acetilado (Yokogawa *et al.*, 1974). Los factores que tienen especial incidencia en la especificidad de las lisinas son el contenido y la naturaleza de los ácidos teicóicos (TAs) y lipoteicóicos (LTAs) que decoran la pared bacteriana. En este sentido las lisinas modulares ajustan su especificidad mediante el módulo de anclaje y/o la disposición modular con el módulo catalítico.

Cpl-7 es capaz de admitir y estabilizar en el centro catalítico las cadenas glicánicas y peptídicas del peptidoglicano, para su hidrólisis. Este centro catalítico está constituido por aminoácidos conservados en la familia GH-25 y además tiene una alta homología con Cpl-1. Como se ha indicado a lo largo de esta memoria, no se conoce la naturaleza del grupo o grupos que actúan como receptores del módulo de unión a la pared de Cpl-7, pero si se sabe que la presencia del módulo de unión a la pared es esencial para su actividad y que además es independiente de colina (Díaz *et al.*, 1991; Sanz *et al.*, 1992a).

Dentro de los distintos ensayos realizados para identificar dichos elementos reconocidos por Cpl-7, se intentó la obtención de complejos cristalográficos con varios compuestos estructuralmente relacionados con los TAs y el peptidoglicano, como el ribitol, ribitol-fosfato, TAs purificados de neumococo y con fragmentos sintéticos de la pared (disacáridos, monosacáridos-monopéptido y disacárido-monopéptido). En ningún caso se obtuvieron complejos, quizá, porque no tienen una gran afinidad por el módulo de anclaje o quizá, porque el bajo contenido de solvente en los cristales y/o el empaquetamiento del cristal no permitieron su difusión a través de los canales de solvente.

La composición de los ácidos teicóicos varía significativamente de unas bacterias a otras. Frente a la complejidad que muestra neumococo, pueden llevar unido D-Ala y GalNAc (Draing *et al.*, 2006), entre muchos otros componentes. Los TAs de *S. pyogenes* o *S. agalactiae*, también pueden ser reconocidos por fago que codifican para endolisinas portadoras de repeticiones que conforman el CWBM de Cpl-7 (Beres *et al.*, 2006; Donovan & Foster-Frey, 2008). Estos TAs son simples polímeros de glicerolfosfato que pueden estar modificados por la adición de D-Ala y glucosa a los residuos de glicerol (Goldschmidt & Panos, 1984; Scott & Barnett, 2006; Weidenmaier & Peschel, 2008). A pesar de que los residuos de D-Ala son un elemento común a todos ellos, la posibilidad de que actúen como receptores para los motivos CWBM de Cpl-7 puede descartarse, ya que los TAs de la cepa R6 de neumococo carecen de esta modificación y Cpl-7 es, sin embargo, capaz de hidrolizarla (Kovacs *et al.*, 2006). Esto deja como elementos comunes, o con una cierta similitud estructural entre los distintos TAs, a los grupos fosfatos y al glicerol/ribitol. En principio parece poco probable que los grupos fosfatos sean reconocidos por Cpl-7 ya que en la superficie del CWBM predominan los residuos ácidos, y carece de los residuos positivos necesarios como para establecer la unión a grupos fosfato. Así mientras que proteínas que unen ADN y que presentan motivos estructuralmente parecidos a Cpl-7, como es el caso de un regulador de la transcripción (Tahirov *et al.*, 2002) (código PDB:1h88), poseen una superficie de potencial electrostático positiva (figura 4.18-B), esto no ocurre en el CWBM de Cpl-7 cuyo carácter es esencialmente ácido (figura 4.16-B), lo que debe de impedir la interacción con grupos cargados negativamente como los fosfatos.

5.1.6 Disposición modular de la endolisina Cpl-7

Como se ha puesto de manifiesto en trabajos previos, la agrupación de los distintos módulos en las endolisinas juega un papel fundamental, no solamente en la capacidad de fijar la enzima sobre sustratos tan complejos como la pared celular; es decir no sólo aumentando la afinidad de la

endolisina; sino también en algunos casos regulando su especificidad. Así ocurre también, en enzimas modulares de neumococo como la autolisina LytC cuya disposición modular provoca que la enzima se encuentre inactiva ante el peptidoglicano intacto de modo que la activación de la enzima requiere la hidrólisis previa por una amidasa de las conexiones peptídicas entre cadenas glicánicas (Pérez-Dorado *et al.*, 2010). En este sentido el aumento de la afinidad por el sustrato requiere de la presencia de módulos especializados en el reconocimiento a la pared, pero no de una interacción marcada entre los módulos; este sería el caso de la endolisina PlyM. Sin embargo, el aumento de la especificidad va asociado con la presencia de interacciones fuertes entre módulos mediada por interacciones hidrofóbicas y polares que permiten al tiempo una gran plasticidad entre el dominio catalítico y el de anclaje, como ocurre en el caso de Cpl-1 (Martínez-Buey *et al.*, 2007).

En el mencionado caso de Cpl-1, la interacción modular se lleva a cabo mediante una cavidad hidrofóbica en el módulo catalítico que se complementa con el extremo hidrofóbico del módulo C-terminal. El conector entre ellos tiene 10 aminoácidos y no establece interacciones con ninguno de los dos módulos debido a la flexibilidad y carga negativa de sus aminoácidos. Aunque, el conocimiento de la disposición modular de Cpl-7 requeriría la determinación de la estructura cristalina de la proteína completa, sin embargo el estudio estructural de los módulos por separado nos proporciona información importante de cómo sería su plegamiento *in vivo*.

La cavidad hidrofóbica encontrada en el módulo catalítico, responsable de estabilizarlo respecto al módulo de anclaje en Cpl-1 está totalmente conservada en Cpl-7, pero no en otros miembros de la familia GH-25. Por otro lado atendiendo a su composición, el conector entre ambos módulos en Cpl-7 es más largo (16 aminoácidos) que en Cpl-1 (10 aminoácidos). Este conector es homólogo a Cpl-1 en el inicio de su secuencia, donde predominan los residuos ácidos (conector de Cpl-7: **DDEKEDNINNENTLKS**, conector de Cpl-1: **DDEEDDKPKT**). En Cpl-1 este carácter ácido promueve la salida del conector hacia fuera de la superficie del módulo catalítico, por lo que es muy posible que ocurra lo mismo en Cpl-7. Dado que el conector de Cpl-7 es más largo y posiblemente más flexible (carece de prolina). Cpl-7 podría albergar al módulo de anclaje un poco más alejado que en Cpl-1. La presencia de ambos factores, la cavidad hidrofóbica en el módulo catalítico y un conector flexible y ácido son fuertes evidencias para suponer que Cpl-7 tendría una disposición modular semejante a la de Cpl-1, a pesar de que la estructura tridimensional de los módulos de anclaje en ambos casos son muy diferentes.

El estudio estructural de las repeticiones idénticas del módulo de anclaje de Cpl-7, indica cuáles podrían ser las interacciones que estabilizarán los tres dominios repetidos en su módulo de unión a la pared. La identidad entre las repeticiones proporcionaría en Cpl-7 una conformación semicircular y compacta diferente de la conformación alargada que presenta RecX y diferente del modelo de SAXS presentado recientemente (Bustamante *et al.*, 2010). La estructura de Cpl-7 constaría de dos módulos compactos pero con una gran plasticidad estructural gracias a la longitud del conector y a las interacciones hidrofóbicas con su módulo catalítico lo que permitiría una

regulación de su interacción en las diferentes familias de bacterias en las que este módulo está presente.

5.2 IMPLICACIONES BIOLÓGICAS DEL RECEPTOR SIGN-R1

En este estudio se han determinado las estructuras tridimensionales del dominio de reconocimiento a carbohidratos de SIGN-R1 y sus complejos con sulfodextrano y ácido siálico. Estas estructuras revelan la configuración del sitio canónico de unión a carbohidratos dependiente de calcio (Ca^{2+}) que está conservado en todos los miembros de la familia de las CLTDs. La estructura de SIGN-R1 no presenta los calcios estructurales Ca^{2+}_2 y Ca^{2+}_3 , descritos para otras CLTDs. Esta ausencia en SIGN-R1 parece asociada a mutaciones puntuales, por residuos de naturaleza hidrofóbica que cambian la conformación del bucle LLR, confiriendo unas características únicas de unión al sustrato a SIGN-R1. En efecto este cambio estructural en SIGN-R1, promueve que residuos incluidos en la estabilización de los calcios estructurales en DC-SIGN(R), estén aquí involucrados en la estabilización del sitio primario de unión a carbohidratos, como se hizo evidente en el estudio de los complejos cristalográficos de SIGN-R1 con sulfodextrano y ácido siálico (figura 4.34 4.35 y 4.37). En ambos casos la Asn258 forma parte de la red de puentes de hidrógeno que estabilizan al ligando en el sitio canónico y no a los calcios como ocurre en DC-SIGN.

5.2.1 SIGN-R1 en la inhibición del proceso inflamatorio

SIGN-R1 cumple un papel fundamental en la inhibición de los procesos inflamatorios como receptor de los anticuerpos glicosilados, reconociendo específicamente los enlaces glicosídicos α -2,6 entre la galactosa y el ácido siálico terminal y no los enlaces glicosídicos α -2,3 (Anthony & Ravetch, 2010). Únicamente se tiene información estructural acerca de un anticuerpo sialilado en α -2,3, (código PDB: 1MCO). Esta estructura indica que el grupo carboxílico del ácido siálico se encuentra apuntando hacia el interior de la proteína y por tanto no estaría expuesto. La estructura obtenida en el complejo SIGN-R1:Siálico indicó que la interacción con el siálico viene mediada a través del grupo carboxílico del ácido siálico (figura 4.37), de modo que, no podría existir la interacción entre SIGN-R1 y un anticuerpo sialilado en α -2,3. Con el fin de aclarar cómo podría ser el reconocimiento de SIGN-R1 con el anticuerpo sialilado en α -2,6, se cristalizó y se determinó la estructura de la porción Fc del anticuerpo sialilado con un enlace glicosídico en α -2,6.

Los datos de difracción obtenidos de la porción Fc tenían una altísima resolución (2.01 Å) y se observó toda la cadena glicánica excepto el ácido siálico terminal. A pesar de ello, la estructura revela que la galactosa previa al siálico, tiene una conformación completamente distinta de la observada en el caso del anticuerpo en donde el ácido siálico se encuentra enlazado α -2,3. La estructura de la galactosa de nuestro anticuerpo provocaría que el grupo carboxílico en nuestra estructura debería estar apuntando hacia fuera de la superficie de la proteína, permitiendo por tanto la interacción con SIGN-R1. Con el fin de comprobar la disponibilidad del grupo carboxílico del ácido siálico se han realizado estudios de dinámica molecular en los anticuerpos sialilados α -2,3 y α -2,6. Estos estudios

confirman que el ácido carboxílico del ácido siálico en posición α -2,6 queda expuesto mientras que cuando el enlace glicosídico es α -2,3 el grupo carboxílico queda enterrado en la superficie de la proteína. De este modo nuestro complejo de SIGN-R1 con el ácido siálico explicaría la especificidad de SIGN-R1 por los anticuerpos sialilados en posición α -2,6 y no en α -2,3. Este tipo de reconocimiento por parte de SIGN-R1 a través del ácido siálico, se podría extrapolar al reconocimiento de otras glicoproteínas del sistema inmune con ácido siálico terminal en posición α -2,6. De esta forma se explicaría el papel fundamental que cumple SIGN-R1 en la inhibición de los procesos inflamatorios, los cuales son inhibidos por un exceso de anticuerpos sialilados α -2,6.

5.2.2 SIGN-R1 en la activación de la vía clásica del complemento

El trabajo estructural ha revelado la presencia de un sitio secundario de unión a carbohidratos, único en SIGN-R1, que no es dependiente de la unión a calcio. Interessantemente, este sitio se sitúa en un canal en la cara opuesta del sitio canónico muy lejos del sitio donde se observa el Ca1. De modo que ambos sitios el primario y el secundario son independientes en términos estructurales (figura 5.1).

La estructura de los cristales nativos y en complejo con sulfodextrano, revelaron la capacidad que posee SIGN-R1 para reconocer a través del sitio secundario patrones moleculares repetitivos como moléculas de sulfato, así como moléculas de glucosas-4-sulfato. Estas moléculas se disponen de forma equidistante a lo largo del canal que conforma el sitio secundario.

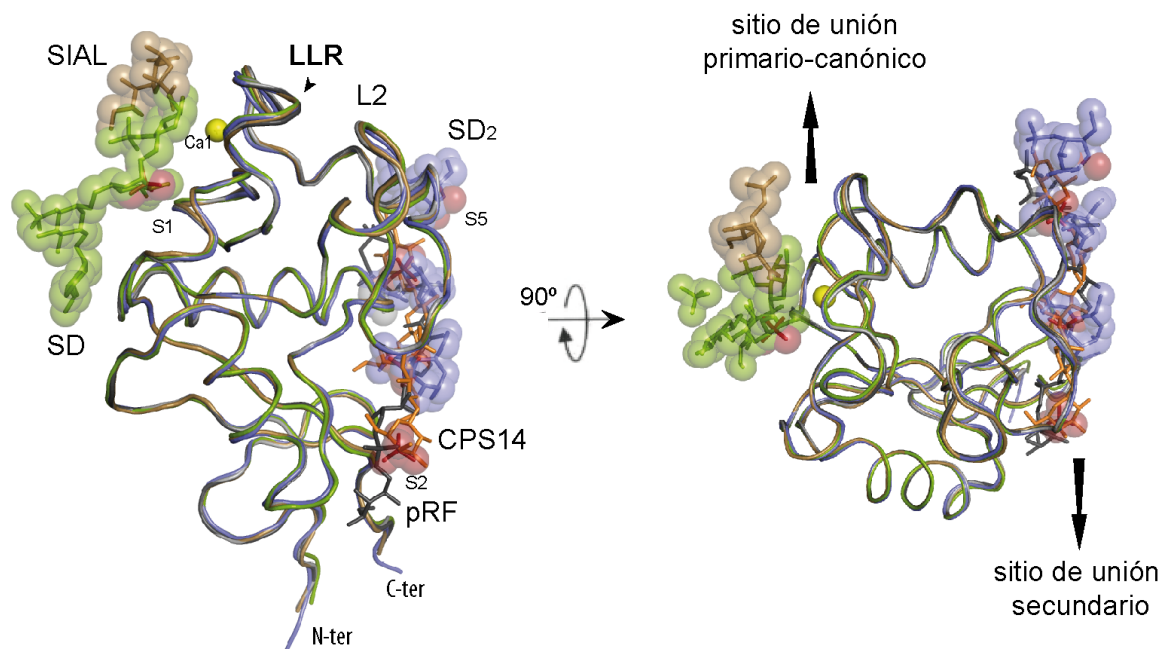


Figura 5.1. Sitios primario y secundario de unión a carbohidratos en SIGN-R1. Se representa una superposición de los ligandos obtenidos en los complejos cristalográficos y los resultados de “docking”. Se puede observar claramente el sitio de unión primario y el secundario, independientes estructuralmente. El sitio secundario se ubica en la cara opuesta del sitio primario (Ca1, esfera en amarillo), orientado al extremo C-terminal. SIGN-R1:Sulfodextrano (SD, en verde y azul); SIGN-R1:Siálico (SIAL, en marrón); resultados de “docking” para el poli-ribitol-fosfato (pRF, en varillas gris) y la cápsula polisacárida 14 de neumococo (CPS14, en varillas marañas), las moléculas de sulfato se representa en esferas de color rojo.

Los ensayos de “docking” apuntan que el poli-ribitol-fosfato (componente de ácidos teicoicos de algunos patógenos) puede ser acomodado en el sitio secundario de unión que hemos encontrado

en SIGN-R1. Los mismos resultados se han observado con el modelo de la cápsula polisacárida del serotipo 14 (pentasacárido) de neumococo, en el que los anillos glicánicos de CPS14 reemplazan los lugares ocupados por los anillos de las glucosas-4-sulfato obtenidos en el complejo de SIGN-R1:Sulfodextrano (figura 4.37 y 4.38).

La existencia de este sitio secundario en SIGN-R1, viene reforzado por el bloqueo del reconocimiento de la CPS14 por parte de los macrófagos, al agregar anticuerpos producidos contra los últimos 15 aminoácidos de este receptor (Kang *et al.*, 2004), donde se localiza el sitio secundario. Este resultado indica la presencia en SIGN-R1 de un sitio de unión distinto al canónico, y descartaría al sitio primario como el único sitio de reconocimiento al patógeno.

Para obtener más pruebas experimentales de la presencia de este nuevo sitio en SIGN-R1 el grupo del Dr. Chae Gyu Park (Laboratory of Cellular Physiology and Immunology in the Chris Browne Center for Immunology and Immune Diseases, Rockefeller University), ha realizado un estudio comparativo entre SIGN-R1 y DC-SIGN. Para ello, tanto SIGN-R1 como DC-SIGN fueron expresadas en células de ovario de hámster, células CHO. Estas células CHO fueron incubadas con dextrano marcado con un cromóforo fluorescente (FITC-dextrano). La fluorescencia de estas células fue medida mediante un clasificador de células activadas por fluorescencia, FACS (de sus siglas en inglés *Fluorescence Activated Cell Sorter*). Para estudiar la dependencia del sitio primario en el reconocimiento del dextrano, se realizaron estos ensayos a concentraciones crecientes de EDTA y sin EDTA, para quelar el calcio del sitio primario, y así bloquear el reconocimiento de este polisacárido por el sitio primario de ambos receptores.

Los resultados obtenidos de este ensayo permitieron concluir que el reconocimiento de ambos receptores por el dextrano es diferente. SIGN-R1 es capaz de unir mayor cantidad de dextrano a concentraciones crecientes de EDTA que DC-SIGN, en concordancia con los dos sitios de unión encontrados en SIGN-R1 frente al único sitio de unión en DC-SIGN. Mientras que DC-SIGN pierde la actividad de forma radical a 3mM de EDTA, SIGN-R1 disminuye la actividad progresivamente, (figura 5.2-A). Vale la pena destacar que la superficie de potencial electrostático y la cavidad en el sitio secundario son diferentes en ambos receptores (figura 5.2-B).

Todos estos resultados indican que SIGN-R1 presenta dos sitios de unión a carbohidrato independientes, el sitio de unión a carbohidrato canónico que depende de las interacciones con el átomo de calcio, y un segundo sitio de baja afinidad que podría reconocer componentes esenciales con patrones repetitivos existentes en la superficie de los patógenos, tales como el dextrano, sulfodextrano, ácidos teicoicos o cápsula polisacárida de neumococo.

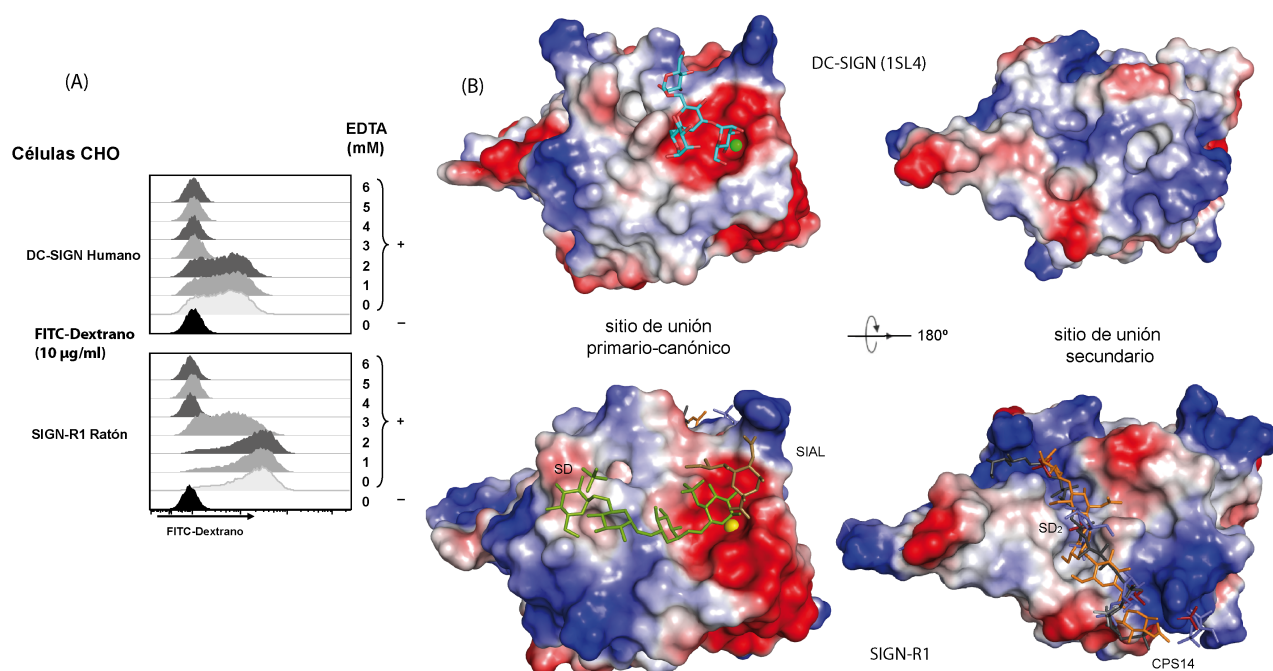


Figura 5.2. Ensayos de unión a FITC-Dextrano a diferentes concentraciones EDTA. Arriba expresión de DC-SIGN en células CHO, abajo expresión de SIGN-R1 en células CHO. El cultivo celular fue incubado con concentraciones decrecientes de EDTA (+) y sin EDTA (-) y el FITC-dextrano se mantuvo constante para todos los ensayos a 10 µg·ml⁻¹, 37 °C, 5% CO₂ durante 1 hora, para un sucesivo análisis de FACS. La línea base de cada ensayo corresponde al control negativo, sin EDTS y sin FITC-Dextrano.

Como ya se ha explicado anteriormente, SIGN-R1 es el receptor crucial en la activación de la vía clásica del complemento a través de la interacción directa con C1q (Kang *et al.*, 2006), pero los mecanismos por los cuales esta activación es llevada a cabo aún se desconocen. C1q es un complejo proteico formado por seis subunidades, y cada subunidad es un heterotrímero (cadenas A, B y C), la cadena A de este heterotrímero está glicosilada en la Asn124 (Gaboriaud *et al.*, 2003). La composición de la cadena glicánica de la C1q ha sido determinada, mostrando exactamente la misma naturaleza que la cadena glicánica que esta presente en los Fc-sialilados, (Mizuochi *et al.*, 1978), en ambos casos el ácido siálico está unido al carbono C-6 de la galactosa terminal, a través de un enlace glicosídico α 2,6. La estructura del complejo SIGN-R1:Siálico ha mostrado que esta interacción se produce a través del sitio primario mediante una interacción directa del Ca1 y el grupo carboxílico del ácido siálico.

Considerando todos estos resultados, se puede proponer un modelo de activación de la vía clásica del complemento mediado por SIGN-R1, en el que SIGN-R1 capturaría al complejo C1 del complemento (C1q-C1s-C1r), a través de una interacción directa entre el sitio de unión primario y el ácido siálico terminal α -2,6 de la cadena polisacárida de la C1q. Además la presencia en SIGN-R1 de un sitio secundario de unión, estructural y funcionalmente independiente del sitio primario, capaz de reconocer motivos repetitivos permite el reconocimiento de motivos de superficie de patógenos, tales como la cápsula polisacárida de neumococo o componentes de la pared como los ácidos teicoicos, (figura 5.3). Este mecanismo le permitiría a SIGN-R1 la identificación de patógenos y al mismo tiempo reclutar la maquinaria para el ensamblado y fijación de C3 con gran eficacia.

De este modo podríamos considerar a SIGN-R1 como un receptor bifuncional capaz, al mismo tiempo, de reconocer específicamente al patógeno y a las proteínas del hospedador. El análisis estructural de SIGN-R1 revela así las bases moleculares de la activación de la vía del complemento por SIGN-R1 y descubre la enorme versatilidad de la familia de lectinas dependientes de Calcio, donde mutaciones puntuales provocan cambios estructurales y funcionales muy importantes que han sido usados por los mamíferos en su lucha constante contra las infecciones de los patógenos bacterianos.

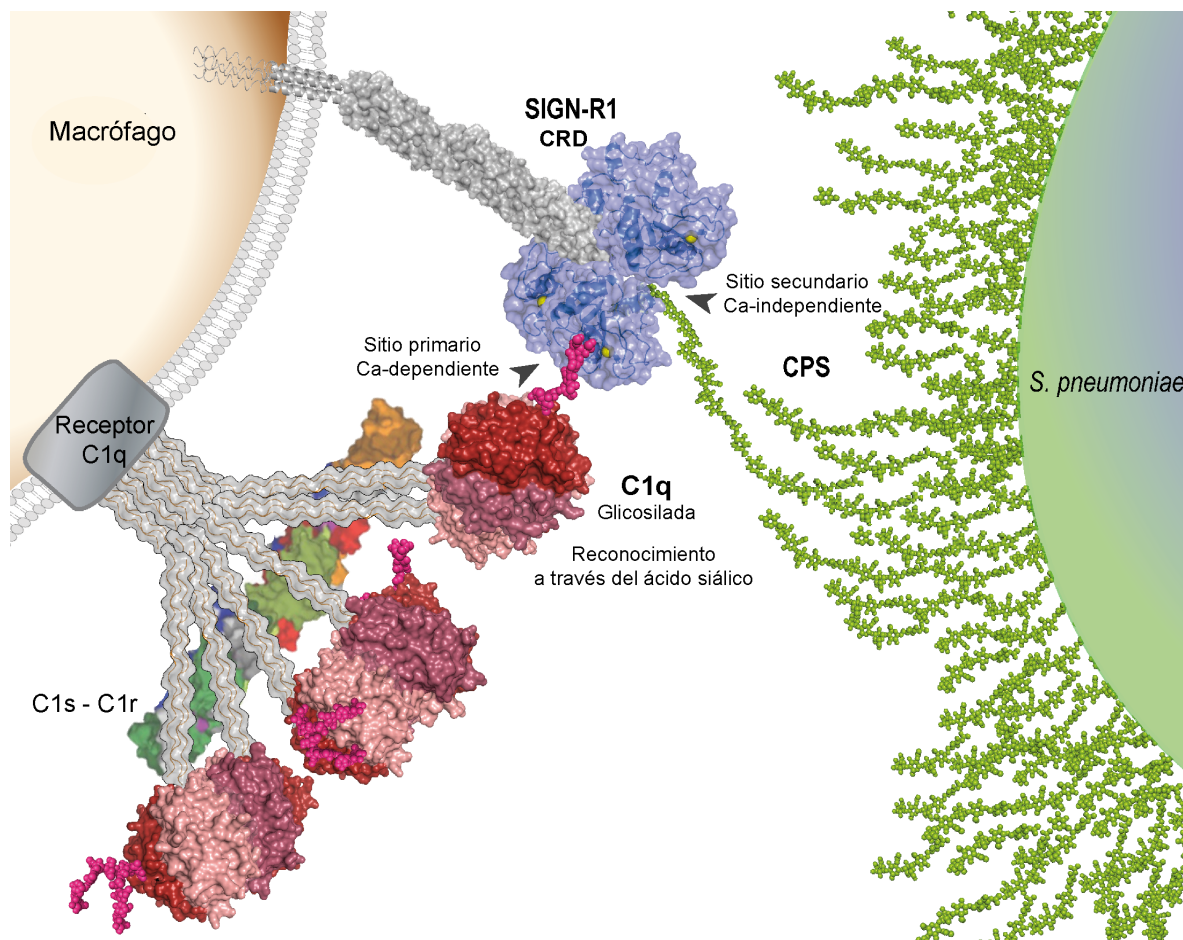


Figura 5.3. Modelo de fijación de la C1q y *S. pneumoniae* por SIGN-R1. El receptor SIGN-R1 emerge de la membrana del macrófago como tetrámero, en gris se representa la fracción extracelular con estructura desconocida hasta el momento, en azul se representa el CRD_SIGN-R1 determinado en este trabajo y el único calcio se representa en esferas amarillas, en donde se señala el sitio de unión primario capaz de reconocer la glicosilación (modelada con esferas de color rosa) sobre la cadena A del heterotrímero de la C1q (rojo). Simultáneamente la SIGN-R1 reconoce la CPS14 (modelada con esferas verdes) en el sitio secundario de unión a carbohidratos.

CONCLUSIONES

6 CONCLUSIONES

Las conclusiones del trabajo recogido en esta memoria se pueden concretar en:

1. Se han determinado las estructuras tridimensionales del módulo catalítico y del dominio de anclaje a la pared bacteriana de la endolisina modular Cpl-7 a resoluciones de 2.4 Å y 1.6 Å respectivamente.
2. El análisis estructural ha revelado que el módulo catalítico de Cpl-7 tiene un plegamiento común con las glicosilhidrolasas de la familia 25 y ha mostrado los residuos críticos en la catálisis y en el reconocimiento del PG tanto en su porción glicánica como peptídica. La cavidad catalítica sería capaz de acomodar cinco anillos glicánicos desde la posición -2 hasta la +3, así como las extensiones peptídicas asociadas a los anillos MurNAc en posiciones -1 y +2.
3. Se ha mostrado por primera vez la estructura tridimensional de los motivos de unión a la pared bacteriana de Cpl-7. La estructura presenta un núcleo hidrofóbico que aglutina tres hélices α de distinta longitud. Tres de estos motivos con idéntica secuencia se unen a través de conectores de seis amino ácidos para conformar el módulo de anclaje de Cpl-7. Aunque no se ha obtenido la estructura del módulo completo, la naturaleza de las repeticiones, así como las interacciones observadas en el cristal, apuntan a una agrupación compacta de las distintas repeticiones.
4. Se han determinado las estructuras del dominio de unión a carbohidratos de SIGN-R1 y sus complejos con sulfodextrano y con ácido siálico, revelando el sitio primario y canónico de unión a carbohidratos que comparte con el resto de las lectinas de tipo C.
5. La estructura de SIGN-R1 revela la presencia de un único ion Ca que sería responsable de la unión a azúcares a través del sitio primario conservado en todas las lectinas dependientes de calcio. Mutaciones puntuales en SIGN-R1 provocan cambios estructurales que bloquearían la unión de los Ca estructurales presentes en otras lectinas y que confieren al sitio de unión, características específicas dentro de la familia.
6. Se han identificado los residuos y las bases moleculares del reconocimiento de azúcares a través del sitio primario dependiente de Ca. En particular, se ha mostrado la estabilización de cadenas glicánicas a través del sitio primario y que la interacción con el ácido siálico se produce a través del grupo carboxilo de este azúcar.
7. Estudios estructurales y de dinámica molecular revelan que las cadenas glicánicas de los anticuerpos Fc sialilados en α -2,6 se dispondrían exponiendo al medio su

grupo carboxilo. Mientras los Fc sialilados en α -2,3 provocan que el carboxilo se encuentre ocluido en interacción directa con el anticuerpo.

8. Considerando la estructura del complejo SIGN-R1:Siálico y los resultados con los anticuerpos Fc_sial, proponemos que el reconocimiento específico de SIGN-R1 por anticuerpos Fc sialilados en α -2,6 se debe a la interacción con el grupo carboxilo del ácido siálico. Esta interacción sería clave en la modulación del estado inflamatorio mediada por SIGN-R1.
9. Tanto la estructura de SIGN-R1 como de su complejo con sulfodextrano han mostrado la presencia de un sitio de unión secundario, estructural y funcionalmente independiente del primario y que sería capaz de unir motivos moleculares repetidos como fosfatos o glucosa-sulfato. Estudios de docking indican que este sitio estabilizaría componentes bacterianos como ácidos teicóicos o como la cápsula polisacárida.
10. Considerando todos estos resultados obtenidos para SIGN-R1, se puede proponer un modelo de activación de la vía clásica del complemento, en el que SIGN-R1 capturaría a C1q, a través de una interacción directa entre el sitio de unión primario y el ácido siálico terminal α -2,6 de la cadena polisacarídica de la C1q, simultáneamente SIGN-R1 secuestraría al patógeno, a través del reconocimiento de motivos repetitivos en la superficie de los patógeno por el sitio secundario. Este mecanismo le permitiría a SIGN-R1 la identificación del patógeno y al mismo tiempo reclutar la maquinaria para el ensamblado y fijación de C3 con gran eficacia, para la eliminación de patógenos encapsulados como el neumococo.

BIBLIOGRAFÍA

7 BIBLIOGRAFÍA

- Abagyan R & Totrov M (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J Mol Biol* **235**, 983-1002.
- Abeyta M, Hardy GG & Yother J (2003) Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of *Streptococcus pneumoniae*. *Infect Immun* **71**, 218-225.
- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC & Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.
- Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403-410.
- Anderson CL, Shen L, Eicher DM, Wewers MD & Gill JK (1990) Phagocytosis mediated by three distinct Fc gamma receptor classes on human leukocytes. *J Exp Med* **171**, 1333-1345.
- Anegón I, Cuturi MC, Trinchieri G & Perussia B (1988) Interaction of Fc receptor (CD16) ligands induces transcription of interleukin 2 receptor (CD25) and lymphokine genes and expression of their products in human natural killer cells. *J Exp Med* **167**, 452-472.
- Anthony RM, Nimmerjahn F, Ashline DJ, Reinhold VN, Paulson JC & Ravetch JV (2008a) Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. *Science* **320**, 373-376.
- Anthony RM & Ravetch JV (2010) A novel role for the IgG Fc glycan: the anti-inflammatory activity of sialylated IgG Fcs. *J Clin Immunol* **30 Suppl 1**, S9-14.
- Anthony RM, Wermeling F, Karlsson MC & Ravetch JV (2008b) Identification of a receptor required for the anti-inflammatory activity of IVIG. *Proc Natl Acad Sci U S A* **105**, 19571-19578.
- Atkinson AP, Cedzynski M, Szemraj J, St Swierzko A, Bak-Romaniszyn L, Banasik M, Zeman K, Matsushita M, Turner ML & Kilpatrick DC (2004) L-ficolin in children with recurrent respiratory infections. *Clin Exp Immunol* **138**, 517-520.
- Austrian R (1953) Morphologic variation in pneumococcus. I. An analysis of the bases for morphologic variation in pneumococcus and description of a hitherto undefined morphologic variant. *J Exp Med* **98**, 21-34.
- Avrameas A, McIlroy D, Hosmalin A, Autran B, Debre P, Monsigny M, Roche AC & Midoux P (1996) Expression of a mannose/fucose membrane lectin on human dendritic cells. *Eur J Immunol* **26**, 394-400.

- Bera A, Biswas R, Herbert S & Gotz F (2006) The presence of peptidoglycan O-acetyltransferase in various staphylococcal species correlates with lysozyme resistance and pathogenicity. *Infect Immun* **74**, 4598-4604.
- Beres SB, Richter EW, Nagiec MJ, Sumby P, Porcella SF, DeLeo FR & Musser JM (2006) Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen group A *Streptococcus*. *Proc Natl Acad Sci U S A* **103**, 7059-7064.
- Bergmann S & Hammerschmidt S (2006) Versatility of pneumococcal surface proteins. *Microbiology* **152**, 295-303.
- Berry AM & Paton JC (2000) Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. *Infect Immun* **68**, 133-140.
- Beveridge TJ (1981) Ultrastructure, chemistry, and function of the bacterial wall. *Int Rev Cytol* **72**, 229-317.
- Borysowski J, Weber-Dabrowska B & Gorski A (2005) [The potential use of bacteriophages in view of the current antibiotic therapy crisis]. *Pol Arch Med Wewn* **113**, 73-78.
- Brändén C-I & Tooze J (1991) *Introduction to protein structure*. New York: Garland Pub.
- Brown JS, Hussell T, Gilliland SM, Holden DW, Paton JC, Ehrenstein MR, Walport MJ & Botto M (2002) The classical pathway is the dominant complement pathway required for innate immunity to *Streptococcus pneumoniae* infection in mice. *Proc Natl Acad Sci U S A* **99**, 16969-16974.
- Brown S, Meredith T, Swoboda J & Walker S (2010) *Staphylococcus aureus* and *Bacillus subtilis* W23 make polyribitol wall teichoic acids using different enzymatic pathways. *Chem Biol* **17**, 1101-1110.
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T & Warren GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921.
- Bustamante N, Campillo NE, García E, Gallego C, Pera B, Diakun GP, Saiz JL, García P, Díaz JF & Menendez M (2010) Cpl-7, a lysozyme encoded by a pneumococcal bacteriophage with a novel cell wall-binding motif. *J Biol Chem* **285**, 33184-33196.
- Cartwright K (2002) Pneumococcal disease in western Europe: burden of disease, antibiotic resistance and management. *Eur J Pediatr* **161**, 188-195.
- Chan AC, Desai DM & Weiss A (1994) The role of protein tyrosine kinases and protein tyrosine phosphatases in T cell antigen receptor signal transduction. *Annu Rev Immunol* **12**, 555-592.
- Chatwell L, Holla A, Kaufer BB & Skerra A (2008) The carbohydrate recognition domain of Langerin reveals high structural similarity with the one of DC-SIGN but an additional, calcium-independent sugar-binding site. *Mol Immunol* **45**, 1981-1994.

- Cleveland RF, Holtje JV, Wicken AJ, Tomasz A, Daneo-Moore L & Shockman GD (1975) Inhibition of bacterial wall lysins by lipoteichoic acids and related compounds. *Biochem Biophys Res Commun* **67**, 1128-1135.
- Collaborative Computational Project N (1994) *Acta Cryst*, 760-763.
- Cundell DR, Gerard NP, Gerard C, Idanpaan-Heikkila I & Tuomanen EI (1995) Streptococcus pneumoniae anchor to activated human cells by the receptor for platelet-activating factor. *Nature* **377**, 435-438.
- d'Herelle F (1931) Bacteriophage as a Treatment in Acute Medical and Surgical Infections. *Bull N Y Acad Med* **7**, 329-348.
- Darden TA & Pedersen LG (1993) Molecular modeling: an experimental tool. *Environ Health Perspect* **101**, 410-412.
- Dawson MH (1933) A Comparative Study of Subcutaneous Nodules in Rheumatic Fever and Rheumatoid Arthritis. *J Exp Med* **57**, 845-858.
- Deresinski S (2009) Bacteriophage therapy: exploiting smaller fleas. *Clin Infect Dis* **48**, 1096-1101.
- Díaz E, López R & García JL (1991) Chimeric pneumococcal cell wall lytic enzymes reveal important physiological and evolutionary traits. *J Biol Chem* **266**, 5464-5471.
- Djurkovic S, Loeffler JM & Fischetti VA (2005) Synergistic killing of Streptococcus pneumoniae with the bacteriophage lytic enzyme Cpl-1 and penicillin or gentamicin depends on the level of penicillin resistance. *Antimicrob Agents Chemother* **49**, 1225-1228.
- Donovan DM & Foster-Frey J (2008) LambdaSa2 prophage endolysin requires Cpl-7-binding domains and amidase-5 domain for antimicrobial lysis of streptococci. *FEMS Microbiol Lett* **287**, 22-33.
- Draing C, Pfitzenmaier M, Zummo S, Mancuso G, Geyer A, Hartung T & von Aulock S (2006) Comparison of lipoteichoic acid from different serotypes of Streptococcus pneumoniae. *J Biol Chem* **281**, 33849-33859.
- Drickamer K (1999) C-type lectin-like domains. *Curr Opin Struct Biol* **9**, 585-590.
- Drickamer K & Taylor ME (1993) Biology of animal lectins. *Annu Rev Cell Biol* **9**, 237-264.
- Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.
- Entenza JM, Loeffler JM, Grandgirard D, Fischetti VA & Moreillon P (2005) Therapeutic effects of bacteriophage Cpl-1 lysin against Streptococcus pneumoniae endocarditis in rats. *Antimicrob Agents Chemother* **49**, 4789-4792.
- Fanger MW, Shen L, Graziano RF & Guyre PM (1989) Cytotoxicity mediated by human Fc receptors for IgG. *Immunol Today* **10**, 92-99.
- Fearon DT & Wong WW (1983) Complement ligand-receptor interactions that mediate biological responses. *Annu Rev Immunol* **1**, 243-271.

- Feinberg H, Guo Y, Mitchell DA, Drickamer K & Weis WI (2005) Extended neck regions stabilize tetramers of the receptors DC-SIGN and DC-SIGNR. *J Biol Chem* **280**, 1327-1335.
- Feinberg H, Mitchell DA, Drickamer K & Weis WI (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* **294**, 2163-2166.
- Felch JW, Inagami T & Hash JH (1975) The N, O-diacetylmuramidase of *Chalaropsis* species. V. The complete amino acid sequence. *J Biol Chem* **250**, 3713-3720.
- Fischer H & Tomasz A (1984) Production and release of peptidoglycan and wall teichoic acid polymers in pneumococci treated with beta-lactam antibiotics. *J Bacteriol* **157**, 507-513.
- Fischer H & Tomasz A (1985) Peptidoglycan cross-linking and teichoic acid attachment in *Streptococcus pneumoniae*. *J Bacteriol* **163**, 46-54.
- Fischer W (1993) Molecular analysis of lipid macroamphiphiles by hydrophobic interaction chromatography, exemplified with lipoteichoic acids. *Anal Biochem* **208**, 49-56.
- Fischer W, Behr T, Hartmann R, Peter-Katalinic J & Egge H (1993) Teichoic acid and lipoteichoic acid of *Streptococcus pneumoniae* possess identical chain structures. A reinvestigation of teichoid acid (C polysaccharide). *Eur J Biochem* **215**, 851-857.
- Fischetti VA (2006) Using phage lytic enzymes to control pathogenic bacteria. *BMC Oral Health* **6 Suppl 1**, S16.
- Fischetti VA, Nelson D & Schuch R (2006) Reinventing phage therapy: are the parts greater than the sum? *Nat Biotechnol* **24**, 1508-1511.
- Fridman WH (1993) Regulation of B-cell activation and antigen presentation by Fc receptors. *Curr Opin Immunol* **5**, 355-360.
- Gaboriaud C, Juanhuix J, Gruez A, Lacroix M, Darnault C, Pignol D, Verger D, Fontecilla-Camps JC & Arlaud GJ (2003) The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. *J Biol Chem* **278**, 46974-46982.
- García E, García JL, García P, Arraras A, Sánchez-Puelles JM & López R (1988) Molecular evolution of lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *Proc Natl Acad Sci U S A* **85**, 914-918.
- García P, García JL, García E, Sánchez-Puelles JM & López R (1990) Modular organization of the lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *Gene* **86**, 81-88.
- García P, Paz Gonzalez M, García E, García JL & López R (1999) The molecular characterization of the first autolytic lysozyme of *Streptococcus pneumoniae* reveals evolutionary mobile domains. *Mol Microbiol* **33**, 128-138.
- García-Bustos JF, Chait BT & Tomasz A (1987) Structure of the peptide network of pneumococcal peptidoglycan. *J Biol Chem* **262**, 15400-15405.

- García-Bustos JF & Tomasz A (1987) Teichoic acid-containing muropeptides from *Streptococcus pneumoniae* as substrates for the pneumococcal autolysin. *J Bacteriol* **169**, 447-453.
- Geijtenbeek TB, Groot PC, Nolte MA, van Vliet SJ, Gangaram-Panday ST, van Duijnhoven GC, Kraal G, van Oosterhout AJ & van Kooyk Y (2002) Marginal zone macrophages express a murine homologue of DC-SIGN that captures blood-borne antigens in vivo. *Blood* **100**, 2908-2916.
- Geijtenbeek TB, Krooshoop DJ, Bleijs DA, van Vliet SJ, van Duijnhoven GC, Grabovsky V, Alon R, Figdor CG & van Kooyk Y (2000) DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking. *Nat Immunol* **1**, 353-357.
- Giesbrecht P, Wecke J & Reinicke B (1976) On the morphogenesis of the cell wall of staphylococci. *Int Rev Cytol* **44**, 225-318.
- Goldschmidt JC, Jr. & Panos C (1984) Teichoic acids of *Streptococcus agalactiae*: chemistry, cytotoxicity, and effect on bacterial adherence to human cells in tissue culture. *Infect Immun* **43**, 670-677.
- Gordon S (2002) Pattern recognition receptors: doubling up for the innate immune response. *Cell* **111**, 927-930.
- Guo Y, Feinberg H, Conroy E, Mitchell DA, Alvarez R, Blixt O, Taylor ME, Weis WI & Drickamer K (2004) Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol* **11**, 591-598.
- Guyre CA, Barreda ME, Swink SL & Fanger MW (2001) Colocalization of Fc gamma RI-targeted antigen with class I MHC: implications for antigen processing. *J Immunol* **166**, 2469-2478.
- Hadfield AT, Harvey DJ, Archer DB, MacKenzie DA, Jeenes DJ, Radford SE, Lowe G, Dobson CM & Johnson LN (1994) Crystal structure of the mutant D52S hen egg white lysozyme with an oligosaccharide product. *J Mol Biol* **243**, 856-872.
- Halgren TA (1995) Potential energy functions. *Curr Opin Struct Biol* **5**, 205-210.
- Hassell AM, An G, Bledsoe RK, Bynum JM, Carter HL, 3rd, Deng SJ, Gampe RT, Grisard TE, Madauss KP, Nolte RT, Rocque WJ, Wang L, Weaver KL, Williams SP, Wisely GB, Xu R & Shewchuk LM (2007) Crystallization of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **63**, 72-79.
- Henrissat B & Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol* **7**, 637-644.
- Hermoso JA, García JL & García P (2007) Taking aim on bacterial pathogens: from phage therapy to enzybiotics. *Curr Opin Microbiol* **10**, 461-472.
- Hermoso JA, Monterroso B, Albert A, Galan B, Ahrazem O, García P, Martínez-Ripoll M, García JL & Menéndez M (2003) Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. *Structure* **11**, 1239-1249.

- Hess B, Bekker H, Berendsen H & Fraaije J (1997) LINCS: a linear constraint solver for molecular simulations. . *J Comput Chem* 1463–1472.
- Holtje JV & Tuomanen EI (1991) The murein hydrolases of *Escherichia coli*: properties, functions and impact on the course of infections in vivo. *J Gen Microbiol* **137**, 441-454.
- Hornak V, Abel R, Okur A, Strockbine B, Roitberg A & Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712-725.
- Huber R, Deisenhofer J, Colman PM, Matsushima M & Palm W (1976) Crystallographic structure studies of an IgG molecule and an Fc fragment. *Nature* **264**, 415-420.
- Hummell DS, Swift AJ, Tomasz A & Winkelstein JA (1985) Activation of the alternative complement pathway by pneumococcal lipoteichoic acid. *Infect Immun* **47**, 384-387.
- Humphrey JH & Grennan D (1981) Different macrophage populations distinguished by means of fluorescent polysaccharides. Recognition and properties of marginal-zone macrophages. *Eur J Immunol* **11**, 221-228.
- Ikeda K, Sannoh T, Kawasaki N, Kawasaki T & Yamashina I (1987) Serum lectin with known structure activates complement through the classical pathway. *J Biol Chem* **262**, 7451-7454.
- Jado I, López R, García E, Fenoll A, Casal J & García P (2003) Phage lytic enzymes as therapy for antibiotic-resistant *Streptococcus pneumoniae* infection in a murine sepsis model. *J Antimicrob Chemother* **52**, 967-973.
- Jedrzejewski MJ (2001) Pneumococcal virulence factors: structure and function. *Microbiol Mol Biol Rev* **65**, 187-207 ; first page, table of contents.
- Jones C (2005) Vaccines based on the cell surface carbohydrates of pathogenic bacteria. *Anais da Academia Brasileira de Ciências* **77**, 293-324.
- Jones TA, Zou JY, Cowan SW & Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**, 110-119.
- Jonsson S, Musher DM, Chapman A, Goree A & Lawrence EC (1985) Phagocytosis and killing of common bacterial pathogens of the lung by human alveolar macrophages. *J Infect Dis* **152**, 4-13.
- Kadioglu A, Weiser JN, Paton JC & Andrew PW (2008) The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* **6**, 288-301.
- Kang YS, Do Y, Lee HK, Park SH, Cheong C, Lynch RM, Loeffler JM, Steinman RM & Park CG (2006) A dominant complement fixation pathway for pneumococcal polysaccharides initiated by SIGN-R1 interacting with C1q. *Cell* **125**, 47-58.

- Kang YS, Kim JY, Bruening SA, Pack M, Charalambous A, Pritsker A, Moran TM, Loeffler JM, Steinman RM & Park CG (2004) The C-type lectin SIGN-R1 mediates uptake of the capsular polysaccharide of *Streptococcus pneumoniae* in the marginal zone of mouse spleen. *Proc Natl Acad Sci U S A* **101**, 215-220.
- Kang YS, Yamazaki S, Iyoda T, Pack M, Bruening SA, Kim JY, Takahara K, Inaba K, Steinman RM & Park CG (2003) SIGN-R1, a novel C-type lectin expressed by marginal zone macrophages in spleen, mediates uptake of the polysaccharide dextran. *Int Immunol* **15**, 177-186.
- Kim JO & Weiser JN (1998) Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of *Streptococcus pneumoniae*. *J Infect Dis* **177**, 368-377.
- Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL & Woods RJ (2008) GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J Comput Chem* **29**, 622-655.
- Kombo DC, Nemethy G, Gibson KD, Rackovsky S & Scheraga HA (1996) Computer-aided discrimination between active and inactive mutants of the N-terminal domain of the bacteriophage lambda repressor. *J Mol Biol* **256**, 517-532.
- Kovacs M, Halfmann A, Fedtke I, Heintz M, Peschel A, Vollmer W, Hakenbeck R & Bruckner R (2006) A functional dlt operon, encoding proteins required for incorporation of d-alanine in teichoic acids in gram-positive bacteria, confers resistance to cationic antimicrobial peptides in *Streptococcus pneumoniae*. *J Bacteriol* **188**, 5797-5805.
- Krapp S, Mimura Y, Jefferis R, Huber R & Sonderrmann P (2003) Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* **325**, 979-989.
- Krarup A, Sorensen UB, Matsushita M, Jensenius JC & Thiel S (2005) Effect of capsulation of opportunistic pathogenic bacteria on binding of the pattern recognition molecules mannan-binding lectin, L-ficolin, and H-ficolin. *Infect Immun* **73**, 1052-1060.
- Kronborg G & Garred P (2002) Mannose-binding lectin genotype as a risk factor for invasive pneumococcal infection. *Lancet* **360**, 1176.
- Kudva IT, Jelacic S, Tarr PI, Youderian P & Hovde CJ (1999) Biocontrol of *Escherichia coli* O157 with O157-specific bacteriophages. *Appl Environ Microbiol* **65**, 3767-3773.
- Laemmli UK (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**, 680 - 685.
- Lanoue A, Clatworthy MR, Smith P, Green S, Townsend MJ, Jolin HE, Smith KG, Fallon PG & McKenzie AN (2004) SIGN-R1 contributes to protection against lethal pneumococcal infection in mice. *J Exp Med* **200**, 1383-1393.

- Laskowski RA, Moss DS & Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**, 1049-1067.
- Leslie AGW (1992) *Crystallographic Computing 5: From Chemistry to Biology*. Oxford University Press.
- Lichenstein HS, Hastings AE, Langley KE, MenDíaz EA, Rohde MF, Elmore R & Zukowski MM (1990) Cloning and nucleotide sequence of the N-acetylmuramidase M1-encoding gene from *Streptomyces globisporus*. *Gene* **88**, 81-86.
- Loeffler JM, Djurkovic S & Fischetti VA (2003) Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. *Infect Immun* **71**, 6199-6204.
- Loeffler JM, Nelson D & Fischetti VA (2001) Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. *Science* **294**, 2170-2172.
- López R, García E, García P, Ronda C & Tomasz A (1982) Choline-containing bacteriophage receptors in *Streptococcus pneumoniae*. *J Bacteriol* **151**, 1581-1590.
- Mac LC & Kraus MR (1950) Relation of virulence of pneumococcal strains for mice to the quantity of capsular polysaccharide formed in vitro. *J Exp Med* **92**, 1-9.
- Mancheno JM, Tateno H, Goldstein IJ, Martinez-Ripoll M & Hermoso JA (2005) Structural analysis of the *Laetiporus sulphureus* hemolytic pore-forming lectin in complex with sugars. *J Biol Chem* **280**, 17251-17259.
- Mancheño JM & Angulo I (2009) Proteínas de fusión con un dominio lectina de tipo beta-trébol, procedimiento de obtención y sus aplicaciones. *CSIC Inventores* **P200800909**.
- Mancheño JM, Tateno H, Goldstein IJ, Martinez-Ripoll M & Hermoso JA (2005) Structural analysis of the *Laetiporus sulphureus* hemolytic pore-forming lectin in complex with sugars. *J Biol Chem* **280**, 17251-17259.
- Markoishvili K, Tsitlanadze G, Katsarava R, Morris JG, Jr. & Sulakvelidze A (2002) A novel sustained-release matrix based on biodegradable poly(ester amide)s and impregnated with bacteriophages and an antibiotic shows promise in management of infected venous stasis ulcers and other poorly healing wounds. *Int J Dermatol* **41**, 453-458.
- Matthews BW (1968) Solvent content of protein crystals. *J Mol Biol* **33**, 491-497.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.
- McCullers JA, Karlstrom A, Iverson AR, Loeffler JM & Fischetti VA (2007) Novel strategy to prevent otitis media caused by colonizing *Streptococcus pneumoniae*. *PLoS Pathog* **3**, e28.
- McCullers JA & Tuomanen EI (2001) Molecular pathogenesis of pneumococcal pneumonia. *Front Biosci* **6**, D877-889.

- McPherson A (1989) *Preparation and analysis of protein crystals*, 2nd ed. Malabar, Fla.: R.E. Krieger Pub. Co.
- McPherson A & DeLucas L (1999) Crystal-growing in space. *Science* **283**, 1459.
- Meroueh SO, Bencze KZ, Heseck D, Lee M, Fisher JF, Stemmler TL & Mobashery S (2006) Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc Natl Acad Sci U S A* **103**, 4404-4409.
- Meyer-Wentrup F, Cambi A, Adema GJ & Figdor CG (2005) "Sweet talk": closing in on C type lectin signaling. *Immunity* **22**, 399-400.
- Mitchell TJ (2003) The pathogenesis of streptococcal infections: from tooth decay to meningitis. *Nat Rev Microbiol* **1**, 219-230.
- Mitchell TJ, Alexander JE, Morgan PJ & Andrew PW (1997) Molecular analysis of virulence factors of *Streptococcus pneumoniae*. *Soc Appl Bacteriol Symp Ser* **26**, 62S-71S.
- Miyamoto S & Kollman PA (1992) Settle an analytical version of the shake and rattle algorithm for rigid water models. *J Comput Chem*, 952-962.
- Mizuguchi K, Deane CM, Blundell TL, Johnson MS & Overington JP (1998) JOY: protein sequence-structure representation and analysis. *Bioinformatics* **14**, 617-623.
- Mizuochi T, Yonemasu K, Yamashita K & Kobata A (1978) The asparagine-linked sugar chains of subcomponent C1q of the first component of human complement. *J Biol Chem* **253**, 7404-7409.
- Mold C, Edwards KM & Gewurz H (1982) Effect of C-reactive protein on the complement-mediated stimulated of human neutrophils by *Streptococcus pneumoniae* serotypes 3 and 6. *Infect Immun* **37**, 987-992.
- Monzingo AF, Marcotte EM, Hart PJ & Robertus JD (1996) Chitinases, chitosanases, and lysozymes can be divided into procaryotic and eucaryotic families sharing a conserved core. *Nat Struct Biol* **3**, 133-140.
- Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255.
- Nagy T, Simpson P, Williamson MP, Hazlewood GP, Gilbert HJ & Orosz L (1998) All three surface tryptophans in Type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands. *FEBS Lett* **429**, 312-316.
- Neth O, Jack DL, Dodds AW, Holzel H, Klein NJ & Turner MW (2000) Mannose-binding lectin binds to a range of clinically relevant microorganisms and promotes complement deposition. *Infect Immun* **68**, 688-693.
- Nimmerjahn F & Ravetch JV (2007) The antiinflammatory activity of IgG: the intravenous IgG paradox. *J Exp Med* **204**, 11-15.

- Olkhova E, Padan E & Michel H (2007) The influence of protonation states on the dynamics of the NhaA antiporter from *Escherichia coli*. *Biophys J* **92**, 3784-3791.
- Otwinowski Z, Borek D, Majewski W & Minor W (2003) Multiparametric scaling of diffraction intensities. *Acta Crystallogr A* **59**, 228-234.
- Overington J, Johnson MS, Sali A & Blundell TL (1990) Tertiary structural constraints on protein evolutionary diversity: templates, key residues and structure prediction. *Proc Biol Sci* **241**, 132-145.
- Park CG, Takahara K, Umemoto E, Yashima Y, Matsubara K, Matsuda Y, Clausen BE, Inaba K & Steinman RM (2001) Five mouse homologues of the human dendritic cell C-type lectin, DC-SIGN. *Int Immunol* **13**, 1283-1290.
- Patthy L (1991) Exons--original building blocks of proteins? *Bioessays* **13**, 187-192.
- Pehr E (1950) Method for determination of the amino acid sequence in peptides. *Acta Chemica Scandinavica* **4**, 283-293.
- Pérez-Dorado I, Campillo NE, Monterroso B, Heseck D, Lee M, Paez JA, García P, Martínez-Ripoll M, García JL, Mobashery S, Menendez M & Hermoso JA (2007) Elucidation of the molecular recognition of bacterial cell wall by modular pneumococcal phage endolysin CPL-1. *J Biol Chem* **282**, 24990-24999.
- Pérez-Dorado I, Gonzalez A, Morales M, Sanles R, Striker W, Vollmer W, Mobashery S, García JL, Martínez-Ripoll M, García P & Hermoso JA (2010) Insights into pneumococcal fratricide from the crystal structures of the modular killing factor LytC. *Nat Struct Mol Biol* **17**, 576-581.
- Pohlmann S, Soilleux EJ, Baribaud F, Leslie GJ, Morris LS, Trowsdale J, Lee B, Coleman N & Doms RW (2001) DC-SIGNR, a DC-SIGN homologue expressed in endothelial cells, binds to human and simian immunodeficiency viruses and activates infection in trans. *Proc Natl Acad Sci U S A* **98**, 2670-2675.
- Powlesland AS, Ward EM, Sadhu SK, Guo Y, Taylor ME & Drickamer K (2006) Widely divergent biochemical properties of the complete set of mouse DC-SIGN-related proteins. *J Biol Chem* **281**, 20440-20449.
- Pozdnyakova O, Guttormsen HK, Lalani FN, Carroll MC & Kasper DL (2003) Impaired antibody response to group B streptococcal type III capsular polysaccharide in C3- and complement receptor 2-deficient mice. *J Immunol* **170**, 84-90.
- Ragone S, Maman JD, Furnham N & Pellegrini L (2008) Structural basis for inhibition of homologous recombination by the RecX protein. *EMBO J* **27**, 2259-2269.
- Ravetch JV & Kinet JP (1991) Fc receptors. *Annu Rev Immunol* **9**, 457-492.
- Resti M, Moriondo M, Cortimiglia M, Indolfi G, Canessa C, Beccioli L, Bartolini E, de Benedictis FM, de Martino M & Azzari C (2010) Community-acquired bacteremic pneumococcal pneumonia in

- children: diagnosis and serotyping by real-time polymerase chain reaction using blood samples. *Clin Infect Dis* **51**, 1042-1049.
- Riesenfeld-Orn I, Wolpe S, García-Bustos JF, Hoffmann MK & Tuomanen E (1989) Production of interleukin-1 but not tumor necrosis factor by human monocytes stimulated with pneumococcal cell surface components. *Infect Immun* **57**, 1890-1893.
- Robertus JD, Monzingo AF, Marcotte EM & Hart PJ (1998) Structural analysis shows five glycohydrolase families diverged from a common ancestor. *J Exp Zool* **282**, 127-132.
- Rodríguez DD, Grosse C, Himmel S, Gonzalez C, de Ilarduya IM, Becker S, Sheldrick GM & Uson I (2009) Crystallographic ab initio protein structure solution below atomic resolution. *Nat Methods* **6**, 651-653.
- Rodríguez-Cerrato V, García P, Del Prado G, García E, Gracia M, Huelves L, Ponte C, López R & Soriano F (2007a) In vitro interactions of LytA, the major pneumococcal autolysin, with two bacteriophage lytic enzymes (Cpl-1 and Pal), cefotaxime and moxifloxacin against antibiotic-susceptible and -resistant *Streptococcus pneumoniae* strains. *J Antimicrob Chemother* **60**, 1159-1162.
- Rodríguez-Cerrato V, García P, Huelves L, García E, Del Prado G, Gracia M, Ponte C, López R & Soriano F (2007b) Pneumococcal LytA autolysin, a potent therapeutic agent in experimental peritonitis-sepsis caused by highly beta-lactam-resistant *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* **51**, 3371-3373.
- Romero A, López R & García P (1990) Characterization of the pneumococcal bacteriophage HB-3 amidase: cloning and expression in *Escherichia coli*. *J Virol* **64**, 137-142.
- Saludjian P, Prange T, Navaza J, Menez R, Guilloteau JP, Ries-Kautt M & Ducruix A (1992) Structure determination of a dimeric form of erabutoxin-b, crystallized from a thiocyanate solution. *Acta Crystallogr B* **48 (Pt 4)**, 520-531.
- Sánchez-Puelles JM, Sanz JM, García JL & García E (1990) Cloning and expression of gene fragments encoding the choline-binding domain of pneumococcal murein hydrolases. *Gene* **89**, 69-75.
- Sanz JM, Díaz E & García JL (1992a) Studies on the structure and function of the N-terminal domain of the pneumococcal murein hydrolases. *Mol Microbiol* **6**, 921-931.
- Sanz JM, García P & García JL (1992b) Role of Asp-9 and Glu-36 in the active site of the pneumococcal CPL1 lysozyme: an evolutionary perspective of lysozyme mechanism. *Biochemistry* **31**, 8495-8499.
- Saphire EO, Stanfield RL, Crispin MD, Parren PW, Rudd PM, Dwek RA, Burton DR & Wilson IA (2002) Contrasting IgG structures reveal extreme asymmetry and flexibility. *J Mol Biol* **319**, 9-18.
- Schleifer KH & Kandler O (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36**, 407-477.

- Scott JR & Barnett TC (2006) Surface proteins of gram-positive bacteria and how they get there. *Annu Rev Microbiol* **60**, 397-423.
- Seo HS, Cartee RT, Pritchard DG & Nahm MH (2008) A new model of pneumococcal lipoteichoic acid structure resolves biochemical, biosynthetic, and serologic inconsistencies of the current model. *J Bacteriol* **190**, 2379-2387.
- Service RF (2004) Orphan drugs of the future? *Science* **303**, 1798.
- Severin A & Tomasz A (1996) Naturally occurring peptidoglycan variants of *Streptococcus pneumoniae*. *J Bacteriol* **178**, 168-174.
- Sheehan MM, García JL, López R & García P (1997) The lytic enzyme of the pneumococcal phage Dp-1: a chimeric lysin of intergeneric origin. *Mol Microbiol* **25**, 717-725.
- Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* **64**, 112-122.
- Silva-Martin N, Molina R, Angulo I, Mancheno JM, García P & Hermoso JA (2010) Crystallization and preliminary crystallographic analysis of the catalytic module of endolysin from Cp-7, a phage infecting *Streptococcus pneumoniae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**, 670-673.
- Simpson HD & Barras F (1999) Functional analysis of the carbohydrate-binding domains of *Erwinia chrysanthemi* Cel5 (Endoglucanase Z) and an *Escherichia coli* putative chitinase. *J Bacteriol* **181**, 4611-4616.
- Snyder GA, Colonna M & Sun PD (2005) The structure of DC-SIGNR with a portion of its repeat domain lends insights to modeling of the receptor tetramer. *J Mol Biol* **347**, 979-989.
- Switzer RC, 3rd, Merrill CR & Shifrin S (1979) A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal Biochem* **98**, 231-237.
- Tahirov TH, Sato K, Ichikawa-Iwata E, Sasaki M, Inoue-Bungo T, Shiina M, Kimura K, Takata S, Fujikawa A, Morii H, Kumasaka T, Yamamoto M, Ishii S & Ogata K (2002) Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter. *Cell* **108**, 57-70.
- Tenner AJ, Robinson SL & Ezekowitz RA (1995) Mannose binding protein (MBP) enhances mononuclear phagocyte function via a receptor that contains the 126,000 M(r) component of the C1q receptor. *Immunity* **3**, 485-493.
- Tessa Wardlaw EWJ, Matthew Hodge (2006) Pneumonia: the forgotten killer of children.: UNICEF's Division of Communication
- Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS & Vorgias CE (1996) Bacterial chitinase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nat Struct Biol* **3**, 638-648.

- Thompson JD, Higgins DG & Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680.
- Tomasz A (1967) Choline in the cell wall of a bacterium: novel type of polymer-linked choline in *Pneumococcus*. *Science* **157**, 694-697.
- Totrov M & Abagyan R (1997) Flexible protein-ligand docking by global energy optimization in internal coordinates. *Proteins Suppl* **1**, 215-220.
- Tuomanen E (1999) Molecular and cellular biology of pneumococcal infection. *Curr Opin Microbiol* **2**, 35-39.
- Tuomanen E, Liu H, Hengstler B, Zak O & Tomasz A (1985) The induction of meningeal inflammation by components of the pneumococcal cell wall. *J Infect Dis* **151**, 859-868.
- Turner ML (1992) Cell adhesion molecules: a unifying approach to topographic biology. *Biol Rev Camb Philos Soc* **67**, 359-377.
- Twort FW & Twort DN (1921) An Investigation on Influenza. *J Hyg (Lond)* **20**, 85-98.
- Vagin A & Teplyakov A (1997) MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Cryst* **30**, 1022-1025.
- Vaguine AA, Richelle J & Wodak SJ (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **55**, 191-205.
- Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE & Berendsen HJ (2005) GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701-1718.
- Verdonk ML, Cole JC, Hartshorn MJ, Murray CW & Taylor RD (2003) Improved protein-ligand docking using GOLD. *Proteins* **52**, 609-623.
- Vocadlo DJ, Davies GJ, Laine R & Withers SG (2001) Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* **412**, 835-838.
- Vollmer W & Tomasz A (2000) The *pgdA* gene encodes for a peptidoglycan N-acetylglucosamine deacetylase in *Streptococcus pneumoniae*. *J Biol Chem* **275**, 20496-20501.
- Walport MJ (2001a) Complement. First of two parts. *N Engl J Med* **344**, 1058-1066.
- Walport MJ (2001b) Complement. Second of two parts. *N Engl J Med* **344**, 1140-1144.
- Wang Q, Trimbur D, Graham R, Warren RA & Withers SG (1995) Identification of the acid/base catalyst in *Agrobacterium faecalis* beta-glucosidase by kinetic analysis of mutants. *Biochemistry* **34**, 14554-14562.
- Wartha F, Beiter K, Albiger B, Fernebro J, Zychlinsky A, Normark S & Henriques-Normark B (2007) Capsule and D-alanylated lipoteichoic acids protect *Streptococcus pneumoniae* against neutrophil extracellular traps. *Cell Microbiol* **9**, 1162-1171.

- Weaver LH, Grutter MG & Matthews BW (1995) The refined structures of goose lysozyme and its complex with a bound trisaccharide show that the "goose-type" lysozymes lack a catalytic aspartate residue. *J Mol Biol* **245**, 54-68.
- Weidenmaier C & Peschel A (2008) Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat Rev Microbiol* **6**, 276-287.
- Weis WI, Kahn R, Fourme R, Drickamer K & Hendrickson WA (1991) Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science* **254**, 1608-1615.
- Weis WI, Taylor ME & Drickamer K (1998) The C-type lectin superfamily in the immune system. *Immunol Rev* **163**, 19-34.
- Weiser JN, Austrian R, Sreenivasan PK & Masure HR (1994) Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. *Infect Immun* **62**, 2582-2589.
- Yokogawa K, Kawata S, Nishimura S, Ikeda Y & Yoshimura Y (1974) Mutanolysin, bacteriolytic agent for cariogenic Streptococci: partial purification and properties. *Antimicrob Agents Chemother* **6**, 156-165.
- Yother J, Leopold K, White J & Fischer W (1998) Generation and properties of a Streptococcus pneumoniae mutant which does not require choline or analogs for growth. *J Bacteriol* **180**, 2093-2101.
- Young R (1992) Bacteriophage lysis: mechanism and regulation. *Microbiol Rev* **56**, 430-481.
- Zelensky AN & Gready JE (2003) Comparative analysis of structural properties of the C-type-lectin-like domain (CTLCD). *Proteins* **52**, 466-477.

APÉNDICE

8 APÉNDICE

Noella Silva-Martin,^a Rafael Molina,^a Ivan Angulo,^a José M. Mancheño,^a Pedro García^b and Juan A. Hermoso^{a*}

^aGrupo de Cristalografía Macromolecular y Biología Estructural, Instituto Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain, and

^bDepartamento de Microbiología Molecular, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

Correspondence e-mail: xjuan@iqfr.csic.es

Received 3 February 2010

Accepted 22 February 2010

Crystallization and preliminary crystallographic analysis of the catalytic module of endolysin from Cp-7, a phage infecting *Streptococcus pneumoniae*

As part of the life cycle of the pneumococcal phage Cp-7, the endolysin Cpl-7 cleaves the glycosidic β 1,4 bonds between *N*-acetylmuramic acid and *N*-acetylglucosamine in the pneumococcal cell wall, resulting in bacterial lysis. Recombinant Cpl-7 was overexpressed in *Escherichia coli*, purified and crystallized using the vapour-diffusion method at 291 K. Diffraction-quality tetragonal crystals of the catalytic module of Cpl-7 were obtained from a mixture of PEG 3350 and sodium formate. The crystals belonged to space group *I*422, with unit-cell parameters $a = 127.93$, $b = 127.93$, $c = 82.07$ Å. Diffraction data sets were collected to 2.4 Å resolution using a rotating-anode generator.

1. Introduction

Bacteriophages, or phages, are viruses that infect bacteria. Following replication inside its bacterial host, the phage is faced with a problem: it needs to exit the bacterium to disseminate its progeny. To solve this problem, double-stranded DNA (dsDNA) phages have evolved a lytic system to weaken the bacterial cell wall, resulting in bacterial lysis. Phage lytic enzymes, or endolysins, are highly efficient molecules that have been refined over millions of years for this exact purpose (Fischetti, 2005). The *Streptococcus pneumoniae* bacteriophage Cp-7 (*Podoviridae* family) contains the *cpl-7* gene which encodes the lytic protein Cpl-7 (342 amino-acid residues; ~38.4 kDa; Fig. 1) possessing muramidase (lysozyme) activity (García *et al.*, 1990).

The Cpl-7 lysozyme belongs to the GH25 family of glycosyl hydrolases (<http://www.cazy.org/fam/GH25.html>); it binds to the bacterial envelope and cleaves the glycosidic β 1,4 bonds between the *N*-acetylmuramic acid and *N*-acetylglucosamine moieties of the glycan chains, breaking down the peptidoglycan (PG) backbone.

Most *S. pneumoniae* endolysins display a modular structure. In addition to the catalytic module, they possess a choline-binding module (CBM) which facilitates their anchorage to the choline-containing teichoic acids of the *S. pneumoniae* cell wall (García *et al.*, 1988). In turn, this CBM is formed by the repetition of a basic unit of about 20 amino acids, which is also found in multiple tandem copies in a large family of surface proteins (the choline-binding protein family) from *S. pneumoniae* and its bacteriophages. Interestingly, Cpl-7 does

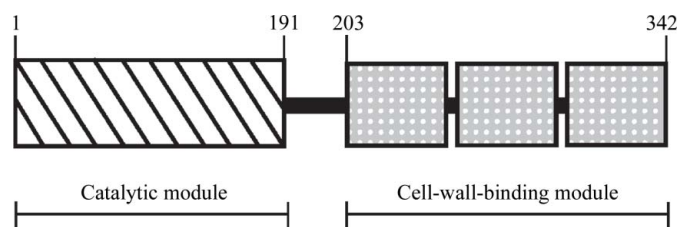


Figure 1
Schematic representation of the modular structure of Cpl-7 endolysin. The catalytic module belongs to glycosyl hydrolase family 25; the linker (black line) and the cell-wall binding module, which is composed of three repeats, are also represented. Residue numbers are indicated at the top.

not present such a choline-binding module, but instead contains a putative cell-wall-binding module formed by three nearly identical tandem repeats of 48 amino acids each (Fig. 1). This cell-wall-binding module presents a completely different sequence to those of other known peptidoglycan-binding motifs. Three complete choline-binding proteins (CBP) have been structurally characterized to date, namely Cpl-1 (Hermoso *et al.*, 2003), Pce (Hermoso *et al.*, 2005) and CbpF (Molina *et al.*, 2007), none of which possess an anchor motif similar to that of Cpl-7.

The Cpl-1 lysozyme (which shares a similar catalytic module to that of Cpl-7) and Pal amidase, which are produced by the Cp-1 and Dp-1 phages, respectively, act in a synergistic manner in a murine model (Jado *et al.*, 2003). Furthermore, using this mouse model it has been demonstrated that Cpl-1 can eliminate colonization caused by *S. pneumoniae* and prevent the development of otitis media (McCullers *et al.*, 2007). These results suggest that these active phage lytic enzymes, termed enzybiotics, have a noticeable effect as antimicrobial agents (Hermoso *et al.*, 2007). Hence, knowledge of the structure of these enzymes will be helpful in the establishment of a novel enzybiotic treatment for streptococci-mediated diseases.

2. Experimental

2.1. Protein expression and purification

The full-length Cpl-7-coding gene was amplified by polymerase chain reaction (PCR) from the pCP700 plasmid using sense (5'-GGA-ATTCATGGTTAAGAAAAATGATTTATTTGTAGACGTTG-3') and antisense (5'-CCCAAGCTTAAATAGCTAGTAGAAAATT-TCTACTAGCTTTTACTTGTGA-3') primers, in which the nucleotides in bold are digestion sites for the restriction enzymes *Eco*RI and *Hind*III, respectively. The resulting purified 1000 kb PCR product was cloned into the expression vector pKLSL_t (Mancheño & Angulo, 2009), rendering plasmid pKLSL_t-Cpl-7, which was used to transform competent *Escherichia coli* BL21 (DE3) cells. The expression vector pKLSL_t is based on pET28a(+) (Novagen, Germany) and incorporates the nucleotide sequence coding for LSL_t, namely the β -trefoil module of the haemolytic lectin LSL_a from the mushroom *Laetiporus sulphureus* (Mancheño *et al.*, 2005) followed by a sequence coding for a linker stretch (ASSS) and a TEV (tobacco etch virus) endoprotease

cleavage site (ENLYFQG; Mancheño & Angulo, 2009). Cloning of the nucleotide sequence coding for LSL_t was performed between the *Nco*I and *Eco*RI restriction sites of pET28a(+). For the production of recombinant LSL_t-Cpl-7 fusion protein, the appropriate recombinant strain was cultured in Luria-Bertani broth at 310 K with aeration; when the culture reached an OD₆₀₀ of about 0.8 protein expression was induced with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) and incubation was continued for 4 h. After centrifugation, the cleared cell lysate was applied onto a Sepharose 4B affinity column (GE Healthcare). After exhaustive washing with 20 mM Tris-HCl pH 8.0 containing 0.1 M NaCl buffer and 0.04%(w/v) sodium azide, the fusion protein LSL_t-Cpl-7 was eluted with 200 mM lactose in 20 mM Tris-HCl pH 8.0 with 0.1 M NaCl and 0.04%(w/v) sodium azide. Fractions containing LSL_t-Cpl-7 were pooled and dialyzed against TEV digestion buffer [50 mM Tris-HCl pH 8.0 with 0.2 M NaCl and 0.04%(w/v) sodium azide]. The dialyzed sample was concentrated using a 10 kDa cutoff protein concentrator (Amicon YM-10) and digested with TEV protease (kindly provided by Roger Williams at MRC, Cambridge) at a protein:TEV ratio of 80:1(w:w). Cpl-7 was separated from the LSL_t fusion tag by size-exclusion chromatography on Superdex 75 (GE Healthcare). All purification steps were performed at 277 K. The enzyme was then concentrated at 277 K using a 10 kDa cutoff protein concentrator (Amicon YM-10). The final protein concentration was determined by UV-Vis absorbance measurements with a Nanodrop ND-1000 spectrophotometer (Thermo Scientific), using a theoretical extinction coefficient of 68 425 M⁻¹ cm⁻¹ at 280 nm.

2.2. Crystallization

A NanoDrop robot (Innovadyne Technologies Inc.) was used to assay initial crystallization conditions using Crystal Screens I, II, Lite, Index and SaltRx from Hampton Research and PACT Suite and JCSG+ Suite from Qiagen. Initial assays were carried out by the sitting-drop vapour-diffusion method at 291 K in Innovaplate SD-2 microplates (Innovadyne Technologies Inc.), mixing 250 nl protein solution with 250 nl precipitant solution and equilibrating against 80 μ l well solution. The stock protein was concentrated to 10 mg ml⁻¹ in 20 mM Tris-HCl pH 8.0 and 0.1 M NaCl.



Figure 2
Crystals of the catalytic module of Cpl-7 endolysin grown at 295 K in 0.2 M sodium formate, 20%(v/v) PEG 3350. The approximate dimensions of the crystals are 0.1 \times 0.01 \times 0.1 mm.

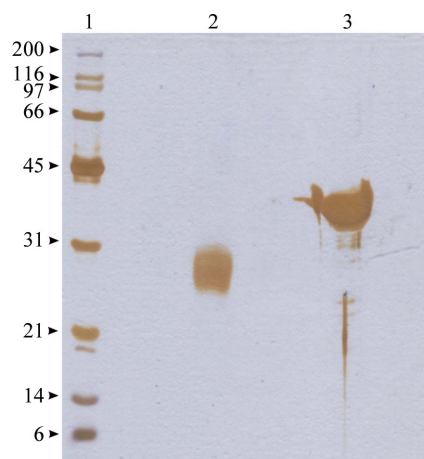


Figure 3
SDS-PAGE analysis of a Cpl-7 crystal. Lane 1, broad-range (BioRad) molecular-weight markers (molecular weights are labelled on the left in kDa); lane 2, dissolved crystal of Cpl-7; lane 3, purified Cpl-7 sample (~38 kDa). The content of the crystal is a partially cleaved Cpl-7 protein compatible with the catalytic module of Cpl-7 (28 kDa). Silver staining was used for protein detection.

2.3. X-ray data collection and processing

X-ray diffraction data were collected at 120 K on an in-house Kappa 2000 Brüker–Nonius CCD detector using Cu $K\alpha$ X-rays generated by an FR591 Brüker–Nonius rotating-anode generator equipped with a double-mirror focusing system and operated at 45 kV and 100 mA. The crystal-to-detector distance was maintained at 200 mm. Crystals were soaked for 10 s in a cryoprotectant solution containing mother liquor and 30% (v/v) glycerol and were then flash-cooled at 100 K in a nitrogen stream using a cryogenic system controlled by a Cryostream Controller 700 (Oxford Cryosystems). The images were processed and scaled using *DENZO* and *SCALEPACK* from the *HKL-2000* suite (Otwinowski & Minor, 1997). Intensities were converted to structure-factor amplitudes using *TRUNCATE* (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

A systematic search for crystallization conditions yielded crystals from 0.2 M sodium formate and 20% PEG 3350 (condition No. 63 of the JCSG+ Suite). Crystals reached maximum dimensions of $0.1 \times 0.01 \times 0.1$ mm in four months (Fig. 2). Because of the long wait for crystals to appear, we decided to run one of the crystals on SDS–PAGE followed by silver staining (Switzer *et al.*, 1979) to assay the crystal composition (Fig. 3). The gel clearly indicated that the protein inside the crystal was partially cleaved, presenting a molecular weight that was compatible with the catalytic module of Cpl-7 (28 kDa). Despite the small size of the crystals, an X-ray data set was collected to 2.4 Å resolution using an in-house rotating-anode source and presented good-quality patterns (Fig. 4). The crystals belonged to the tetragonal system, space group *I*422, with unit-cell parameters $a = b = 127.93$, $c = 82.07$ Å. X-ray data-collection statistics are summarized in Table 1. Specific volume calculations for the catalytic module of Cpl-7 suggested the presence of one molecule in the

Table 1

Data-collection statistics for the Cpl-7 crystal.

Values in parentheses are for the highest resolution shell.

| | |
|---|------------------------|
| Crystal data | |
| Space group | <i>I</i> 422 |
| Unit-cell parameters (Å) | |
| <i>a</i> | 127.93 |
| <i>b</i> | 127.93 |
| <i>c</i> | 83.03 |
| Data processing | |
| Temperature (K) | 100 |
| Wavelength (Å) | 1.5418 |
| Resolution (Å) | 91.28–2.40 (2.49–2.40) |
| Total reflections | 442102 |
| Unique reflections | 37549 |
| Redundancy | 15.40 (12.00) |
| Data completeness (%) | 99.8 (98.50) |
| Average $I/\sigma(I)$ | 20.70 (1.90) |
| Molecules per asymmetric unit | 1 |
| Matthews coefficient (Å ³ Da ^{−1}) | 3.03 |
| Solvent content (%) | 59.50 |
| $R_{\text{merge}}^{\dagger}$ | 0.11 (0.39) |

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

asymmetric unit and a solvent content of 59.50% ($V_M = 3.03$ Å³ Da^{−1}; Matthews, 1968).

Structure solution was initiated with the molecular-replacement method using the structure of the catalytic module of Cpl-1 from bacteriophage Cp-1 (Hermoso *et al.*, 2003; PDB code 1h09) as the initial model. This module shows 84% sequence homology to the catalytic module of Cpl-7 (García *et al.*, 1990). Molecular replacement was performed using the *MOLREP* program (Vagin & Teplyakov, 1998) with reflections in the 15.0–3.5 Å resolution range. A single and unambiguous solution for the rotation and translation functions was obtained, which yielded a final correlation coefficient of 0.64 and an *R* factor of 0.38.

Inspection of the preliminary electron-density map revealed the presence of the complete catalytic module of Cpl-7 but the absence of the cell-wall-binding module. Obviously, the degradation of native Cpl-7 within the crystallization drop, leading to crystallization of the catalytic module, may explain the time that was required for crystal growth. Structural refinement of this module of Cpl-7 is currently in progress.

This work was supported by grants from the Spanish Ministry of Science and Technology (BFU2008-01711, SAF2006-00390), EU-CP223111 (CAREPNEUMO, European Union), the COMBACT program (S-BIO-0260/2006) and CIBER de Enfermedades Respiratorias (CIBERES). CIBERES is an initiative of Instituto de Salud Carlos III. NS-M holds a fellowship from the Fundayacuch Foundation (Venezuela).

References

- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Fischetti, V. A. (2005). *Trends Microbiol.* **13**, 491–496.
- García, E., García, J. L., García, P., Arrarás, A., Sánchez-Puelles, J. M. & López, R. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 914–918.
- García, P., García, J. L., García, E., Sánchez-Puelles, J. M. & López, R. (1990). *Gene*, **86**, 81–88.
- Hermoso, J. A., García, J. L. & García, P. (2007). *Curr. Opin. Microbiol.* **10**, 461–472.
- Hermoso, J. A., Lagartera, L., González, A., Stelter, M., García, P., Martínez-Ripoll, M., García, J. L. & Menéndez, M. (2005). *Nature Struct. Mol. Biol.* **12**, 533–538.

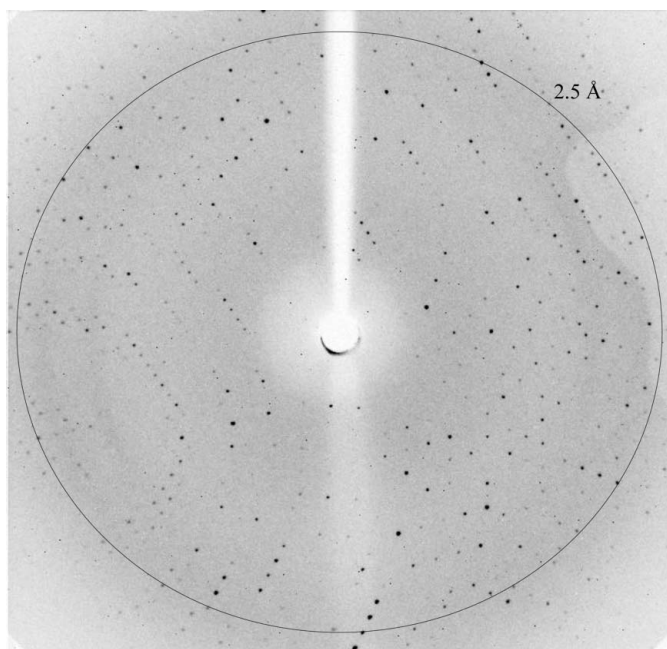


Figure 4
Typical X-ray diffraction pattern of a crystal of the catalytic module of Cpl-7 (oscillation range 1°).

- Hermoso, J. A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P., Martínez-Ripoll, M., García, J. L. & Menéndez, M. (2003). *Structure*, **11**, 1239–1249.
- Jado, I., López, R., García, E., Fenoll, A., Casal, J. & García, P. (2003). *J. Antimicrob. Chemother.* **52**, 967–973.
- Mancheño, J. M. & Angulo, I. (2009). Patent WO/2009/121994.
- Mancheño, J. M., Tateno, H., Goldstein, I. J., Martínez-Ripoll, M. & Hermoso, J. A. (2005). *J. Biol. Chem.* **280**, 17251–17259.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCullers, J. A., Karlstrom, A., Iverson, A. R., Loeffler, J. M. & Fischetti, V. A. (2007). *PLoS Pathog.* **3**, 28.
- Molina, R., González, A., Moscoso, M., García, P., Stelter, M., Kahn, R. & Hermoso, J. A. (2007). *Acta Cryst.* **F63**, 742–745.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Switzer, R. C. III, Merrill, C. R. & Shifrin, S. (1979). *Anal. Biochem.* **98**, 231–237.
- Vagin, A. & Teplyakov, A. (1998). *Acta Cryst.* **D54**, 400–402.

Crystallization and preliminary X-ray diffraction studies of the carbohydrate-recognition domain of SIGN-R1, a receptor for microbial polysaccharides and sialylated antibody on splenic marginal zone macrophages

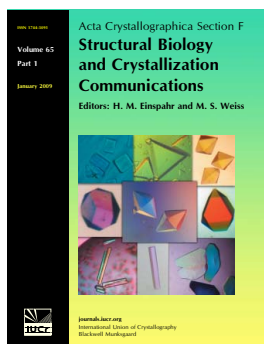
Noella Silva-Martin, Joseph D. Schauer, Chae Gyu Park and Juan A. Hermoso

Acta Cryst. (2009). **F65**, 1264–1266

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



Acta Crystallographica Section F: Structural Biology and Crystallization Communications is a rapid all-electronic journal, which provides a home for short communications on the crystallization and structure of biological macromolecules. It includes four categories of publication: protein structure communications; nucleic acid structure communications; structural genomics communications; and crystallization communications. Structures determined through structural genomics initiatives or from iterative studies such as those used in the pharmaceutical industry are particularly welcomed. *Section F* is essential for all those interested in structural biology including molecular biologists, biochemists, crystallization specialists, structural biologists, biophysicists, pharmacologists and other life scientists.

Crystallography Journals **Online** is available from journals.iucr.org

Crystallization and preliminary X-ray diffraction studies of the carbohydrate-recognition domain of SIGN-R1, a receptor for microbial polysaccharides and sialylated antibody on splenic marginal zone macrophages

Noella Silva-Martin,^a Joseph D. Schauer,^b Chae Gyu Park^b and Juan A. Hermoso^{a*}

^aGrupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química-Física 'Rocasolano', Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain, and ^bLaboratory of Cellular Physiology and Immunology and Chris Browne Center for Immunology and Immune Diseases, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

Correspondence e-mail: xjuan@iqfr.csic.es

Received 2 September 2009

Accepted 13 October 2009

SIGN-R1, or CD209b, is a mouse C-type lectin receptor that is expressed at high levels on macrophages in lymphoid tissues, especially within the marginal zone of the spleen. SIGN-R1 can bind and mediate the uptake of various microbial polysaccharides, including dextrans, lipopolysaccharides and pneumococcal capsular polysaccharides. It has been shown that SIGN-R1 mediates the clearance of encapsulated pneumococcus, complement fixation *via* binding C1q independent of antibody and innate resistance to pneumococcal infection. Recently, SIGN-R1 has also been demonstrated to bind sialylated antibody and mediate its activity to suppress autoimmunity. The carbohydrate-recognition domain (CRD) of SIGN-R1 has been cloned and overexpressed in a soluble secretory form in mammalian Chinese hamster ovary (CHO) cells. The CRD protein of SIGN-R1 was purified from CHO cell-culture supernatant and concentrated for crystallization using the hanging-drop vapour-diffusion method at 291 K. Crystals grew from a mixture of 2 M ammonium sulfate in 0.1 M bistris pH 5.5. Single crystals, which belonged to the monoclinic space group *C*2 with unit-cell parameters $a = 146.72$, $b = 92.77$, $c = 77.06$ Å, $\beta = 121.66^\circ$, allowed the collection of a full X-ray data set to a maximum resolution of 1.87 Å.

1. Introduction

Streptococcus pneumoniae is one of the most common and important human pathogens and causes serious life-threatening diseases such as acute otitis media, pneumonia, sepsis and meningitis. Pneumococcal infections are associated with high morbidity and mortality, especially amongst children, the elderly and immune-depressed patients. The widespread emergence of antibiotic resistance and the lack of highly effective pneumococcal vaccines against all serotypes of this organism give urgency to the elucidation of the molecular processes that are involved in its pathogenicity (Kristinsson, 1997; Pelton, 2000). Recognition of pathogens by the immune system is crucial for the initiation and maintenance of protective immunity. Pattern-recognition receptors, including C-type lectins and Toll-like receptors, discriminate the molecular patterns expressed by pathogens and facilitate differential recognition of pathogens and microbial products (Gordon, 2002).

The innate immune responses provide a critical rapid defence mechanism that acts before the maturation of acquired immunity. Investigations have revealed that SIGN-R1, or CD209b, is a C-type lectin receptor that is primarily found on subsets of macrophages in splenic marginal zone and lymph-node medulla in mouse and mediates the uptake of dextrans (Geijtenbeek *et al.*, 2002; Kang *et al.*, 2003) and lipopolysaccharides (Nagaoka *et al.*, 2005). Furthermore, SIGN-R1 plays an essential role in host defence against *S. pneumoniae* by mediating the recognition of capsular polysaccharide and the clearance of these bacteria (Kang *et al.*, 2004; Park *et al.*, 2009; Lanoue *et al.*, 2004). Moreover, SIGN-R1 contributes to innate resistance by an unusual complement-activation pathway that is independent of immunoglobulin (fulfilling the role of IgM) *via* direct binding with C1q, an essential subcomponent of the classical com-



© 2009 International Union of Crystallography
All rights reserved

plement pathway (Kang *et al.*, 2006). Recently, SIGN-R1 has been identified as a lectin receptor that is required for the anti-inflammatory activity of intravenous sialylated immunoglobulin, suggesting that it is a key player in the new therapeutic approach of using sialylated immunoglobulin to suppress autoimmune diseases (Anthony *et al.*, 2008).

SIGN-R1, a C-type lectin receptor, has a structure consisting of an N-terminal cytosolic domain, a single transmembrane domain, a neck domain and a carbohydrate-recognition domain (CRD) at the C-terminus. We generated and expressed in mammalian cells the soluble form of the CRD part of SIGN-R1 (CRD_SIGN-R1), which contains 134 amino-acid residues. Here, we describe the initial results obtained for the crystallization of CRD_SIGN-R1 and its X-ray diffraction data to 1.87 Å resolution.

2. Experimental procedures

2.1. Expression and purification of CRD_SIGN-R1

To produce a soluble secretory form of CRD_SIGN-R1 protein tagged with a FLAG epitope, the sequences for a signal peptide and a FLAG epitope were fused to the sequence of the CRD portion of SIGN-R1. The construct of soluble FLAG-tagged CRD_SIGN-R1, (SF-CRD_SIGN-R1; GenBank accession No. EU697459) was stably transfected into Chinese hamster ovary (CHO) cells. SF-CRD_SIGN-R1 protein was purified from the culture supernatant of CHO/SF-CRD_SIGN-R1 cells using Anti-FLAG M1 Affinity Gel (Sigma-Aldrich, St Louis, Missouri, USA) following the manufacturer's instructions. 1 mg of protein with a molecular weight of 16.5 kDa was obtained at the end of the purification, which was further concentrated to 3.5 mg ml⁻¹ by centrifugation using an Amicon Ultra-10 membrane in 0.01 M Tris-HCl pH 8 and 0.1 M NaCl buffer.

2.2. Crystallization

Initial assays were carried out by the sitting-drop vapor-diffusion method at 291 K on Innovaplate SD-2 microplates (Innovadyne Technologies Inc.), mixing 250 nl protein solution with 250 nl precipitant solution and equilibrating against 70 µl well solution. High-throughput techniques with a NanoDrop robot (Innovadyne Technologies Inc.) were used to assay crystallization conditions using CRD_SIGN-R1 at 3.5 mg ml⁻¹ in 0.01 M Tris-HCl pH 8 and 0.1 M NaCl with the PACT Suite and JCSG+ Suite from Qiagen and JBScreen Classics 1, 4, 5 and 7 from Jena Bioscience. Successful initial conditions were optimized by hand using hanging-drop methods,



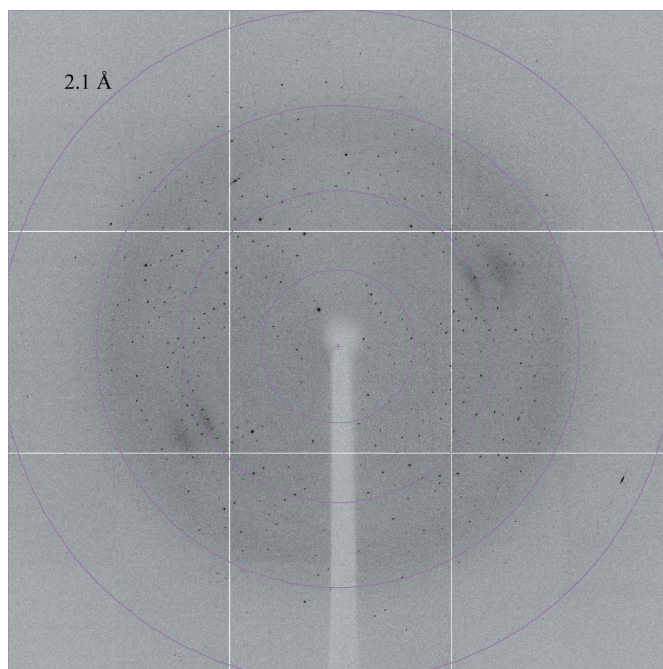
Figure 1

CRD_SIGN-R1 crystals obtained using 1.7 M ammonium sulfate and 0.1 M bis-tris pH 5.5. The approximate dimensions of the crystals were 0.06 × 0.06 × 0.01 mm.

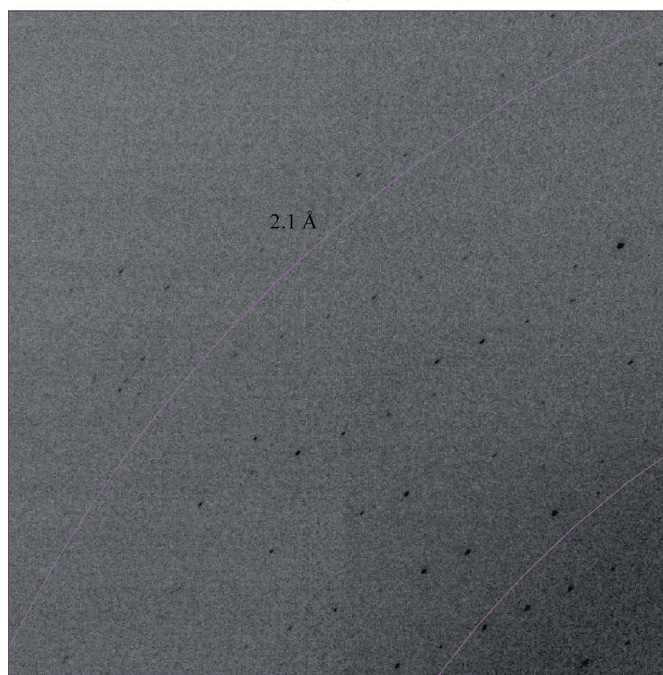
mixing 1 µl protein solution with 1 µl precipitant solution and equilibrating against 500 µl well solution.

2.3. X-ray data collection and processing

All crystals were soaked for 5 s in a cryosalt protective solution consisting of 50%(v/v) saturated lithium sulfate in the crystallization solution prior to flash-cooling to 100 K using a cryogenic system. A native data set was collected using synchrotron radiation on beamline ID23-1 at ESRF (Grenoble) using an ADSC Quantum Q315r



(a)



(b)

Figure 2

(a) X-ray diffraction pattern of CRD_SIGN-R1 crystals (oscillation range 1.0°); (b) enlargement showing the highest resolution area.

Table 1

Data-collection statistics for CRD_SIGN-R1 crystals.

Values in parentheses are for the highest resolution shell.

| | |
|---|------------------------|
| Crystal data | |
| Space group | <i>C</i> 2 |
| Unit-cell parameters | |
| <i>a</i> (Å) | 146.72 |
| <i>b</i> (Å) | 92.77 |
| <i>c</i> (Å) | 77.06 |
| β (°) | 121.66 |
| Data collection | |
| Temperature (K) | 100 |
| Wavelength (Å) | 1.07225 |
| Resolution (Å) | 74.58–1.87 (1.97–1.87) |
| Unique data | 52474 (2478) |
| Multiplicity | 3.9 (2.8) |
| Data completeness (%) | 93.30 (72.54) |
| Average $I/\sigma(I)$ | 11.80 (1.70) |
| Molecules per ASU | 4 |
| Matthews coefficient (Å ³ Da ^{−1}) | 3.49 |
| Solvent content (%) | 64.73 |
| $R_{\text{merge}}^{\dagger}$ | 0.06 (0.57) |

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th measurement of reflection hkl and $\langle I(hkl) \rangle$ is the weighted mean of all measurements.

detector and a wavelength of 1.072250 Å. The collected images were processed and scaled using *MOSFLM* (Leslie, 2006) and *SCALA* (Collaborative Computational Project, Number 4, 1994), respectively.

3. Results

The CRD_SIGN-R1 microcrystals grew under six different conditions. Four of these used ammonium sulfate as the main precipitant agent: (i) 2 *M* ammonium sulfate in 0.1 *M* sodium acetate pH 4.6, (ii) 2 *M* ammonium sulfate with 0.2 *M* sodium chloride in 0.1 *M* sodium cacodylate pH 6.5, (iii) 2 *M* ammonium sulfate in 0.1 *M* Tris–HCl pH 8.5 and (iv) 2 *M* ammonium sulfate in 0.1 *M* bis-tris pH 5.5. The other two conditions used different agents: (v) 2.1 *M* DL-malonic acid pH 7 and (vi) 1.6 *M* trisodium citrate. Good-quality crystals with a rhombohedral shape were obtained using 2 *M* ammonium sulfate in 0.1 *M* bis-tris pH 5.5. The crystals reached maximum dimensions of 0.06 × 0.06 × 0.01 mm in two weeks (Fig. 1).

An X-ray data set was collected to 1.87 Å resolution from a single CRD_SIGNR-1 crystal and displayed apparently weak but well defined diffraction patterns (Fig. 2). X-ray data processing showed good processing statistics (Table 1). The crystal belonged to the monoclinic space group *C*2, with unit-cell parameters $a = 146.72$, $b = 92.77$, $c = 77.06$ Å, $\beta = 121.66^\circ$. Specific volume calculations based on the molecular weight of CRD_SIGNR-1 and the unit-cell parameters indicated the presence of four molecules in the asymmetric unit with 64% solvent content and a Matthews coefficient V_M of 3.49 Å³ Da^{−1} (Matthews, 1968). Structure determination was initiated

using the human DC-SIGN structure (PDB code 1k9j; Feinberg *et al.*, 2001), which shows 68% sequence identity, as the initial structural model. Molecular replacement was performed with the *MOLREP* program (Vagin & Teplyakov, 1997) using reflections to 3.5 Å resolution. Four single and unambiguous solutions for the rotation and translation functions were obtained, which yielded a final correlation coefficient of 0.60 and an *R* factor of 0.49. The space group was confirmed to be *C*2, with four monomers in the asymmetric unit. Structural refinement of the CRD_SIGNR-1 model is currently in progress.

The authors thank the staff of the ID23-1 beamline at ESRF for support during synchrotron data collection. We are also grateful to Dr Ralph Steinman for his support. This work was supported by the Spanish Ministry of Education and Science (BFU2008-01711/BMC and the Factoría de Cristalización from the CONSOLIDER-INGENIO 2010 program). NS-M holds a fellowship from the Fundayacucho Foundation (Venezuela).

References

- Anthony, R. M., Wermeling, F., Karlsson, M. C. & Ravetch, J. V. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 19571–19578.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Feinberg, H., Mitchell, D. A., Drickamer, K. & Weis, W. I. (2001). *Science*, **294**, 2163–2166.
- Geijtenbeek, T. B., Groot, P. C., Nolte, M. A., van Vliet, S. J., Gangaram-Panday, S. T., van Duynhoven, G. C., Kraal, G., van Oosterhout, A. J. & van Kooyk, Y. (2002). *Blood*, **100**, 2908–2916.
- Gordon, S. (2002). *Cell*, **111**, 927–930.
- Kang, Y. S., Do, Y., Lee, H. K., Park, S. H., Cheong, C., Lynch, R. M., Loeffler, J. M., Steinman, R. M. & Park, C. G. (2006). *Cell*, **125**, 47–58.
- Kang, Y. S., Kim, J. Y., Bruening, S. A., Pack, M., Charalambous, A., Pritsker, A., Moran, T. M., Loeffler, J. M., Steinman, R. M. & Park, C. G. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 215–220.
- Kang, Y. S., Yamazaki, S., Iyoda, T., Pack, M., Bruening, S. A., Kim, J. Y., Takahara, K., Inaba, K., Steinman, R. M. & Park, C. G. (2003). *Int. Immunol.* **15**, 177–186.
- Kristinsson, K. G. (1997). *Microb. Drug Resist.* **3**, 117–123.
- Lanoue, A., Clatworthy, M. R., Smith, P., Green, S., Townsend, M. J., Jolin, H. E., Smith, K. G., Fallon, P. G. & McKenzie, A. N. (2004). *J. Exp. Med.* **200**, 1383–1393.
- Leslie, A. G. W. (2006). *Acta Cryst.* **D62**, 48–57.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Nagaoka, K., Takahara, K., Tanaka, K., Yoshida, H., Steinman, R. M., Saitoh, S., Akashi-Takamura, S., Miyake, K., Kang, Y. S., Park, C. G. & Inaba, K. (2005). *Int. Immunol.* **17**, 827–836.
- Park, J. Y., Choi, H. J., Prabagar, M. G., Choi, W. S., Kim, S. J., Cheong, C., Park, C. G., Chin, C. Y. & Kang, Y. S. (2009). *Neurosci. Lett.* **450**, 246–251.
- Pelton, S. I. (2000). *Vaccine*, **19**, Suppl. 1, S96–S99.
- Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

STRUCTURAL BASIS FOR SELECTIVE RECOGNITION OF OLIGOSACCHARIDES BY SIGN-R1, A RECEPTOR FOR MICROBIAL POLYSACCHARIDES AND SIALYLATED ANTIBODY ON SPLENIC MARGINAL ZONE MACROPHAGES

Noella Silva-Martin¹, Sergio G. Bartual¹, Anthony Rodríguez, Chae Gyu Park² & Juan A. Hermoso^{1*}

¹ Department of Crystallography and Structural Biology, Instituto de Química-Física "Rocasolano", CSIC, Serrano 119, 28006-Madrid, Spain. ² Laboratory of Cellular Physiology and Immunology and Chris Browne Center for Immunology and Immune Diseases, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

*Corresponding author email: xjuan@iqfr.csic.es

Summary

The classical complement pathway is the major activation mechanism required for innate protection against pathogen microorganisms. Deposition and activation of C3 convertase over bacterial surface starts the complement cascade leading to the elimination of the microbe. However, encapsulated organisms like *S pneumoniae* are able to resist to the complement pathway limiting the access to cell-bound complement and reducing the amount of C3 deposition.

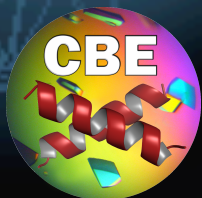
The spleen is the first line defence against blood-borne infections by encapsulated organisms. Recognition and uptake of pneumococcal polysaccharide capsule is carried out principally by C-type lectin like domain family like the human DC-SIGN and the murine SIGN-R1. Recently it was demonstrate that SIGN-R1 could initiate a classical, but Ig-independent complement pathway for C3 fixation, by direct binding of sialilated C1q protein and hence the C2 C4 convertase to produce C3.

That strategy will prevent capsulated pneumococcal cells to scape from C3 opsonisation. Mouse SIGN-R1 is so far the only member of SIGN family that function by binding polysaccharides from endogenous and microbial origins. Furthermore it was reported that SIGN-R1 is able to bind α -2,6-sialilated antibodies supporting the finding of SIGN-R1 to selectively binding α -2,6-syalilated glyco proteins such as C1q.

In attempt to shell light on SIGN-R1 glyco-binding selectivity we have obtained the X-Ray atomic structures of SIGN-R1 globular lectin domain in complex with sulfodextran and sialic acid and the atomic structure of an α -2,6-syalilated antibody coupled with CHO cells binding experiments. In addition molecular dockings with the CSP14 capsule were performed. We found that syalilated glyco-proteins could bind to the classical lectin polysaccharide recognition zone while CSP seems to have a new secondary recognition region. This two binding sites will allow SIGN-R1 to bind sialilated glyco-proteins and CSP simultaneously.



UNIVERSIDAD
AUTÓNOMA
DE MADRID



Departamento de
Cristalografía y
Biología Estructural



Consejo Superior
de Investigaciones
Científicas

